THE INTERNATIONAL WEEKLY JOURNAL OF SCIENCE

nature

THE MEDIA ON SCIENCE

Cheerleader or watchdog?

CELL LONGEVITY
A touch of germline immortality

PLANT HORMONES
Ten and counting

AN OCEAN ON ENCELADUS
The sodium test

NATUREJOBS

Vol 459 | Issue no. 7250 | 25 June 2009

Abstractions

LAST AUTHOR

People with Down's syndrome, a disorder caused by having an extra copy of all or part of chromosome 21, tend to have lower-than-average cognitive ability and several characteristic physical features, as well as a shorter lifespan. But, much to the puzzlement of scientists, they are also less likely to develop solid cancers — tumours that form in tissues or organs. Theories proposed to explain this phenomenon include the possibility that many individuals with Down's syndrome die too early for such cancers to emerge, or that they aren't exposed to environmental risk factors or to stress and pressure, which have been linked to cancer. But Sandra Ryeom at the Children's Hospital Boston in Massachusetts and her team have discovered a biological basis for the protective effect of the disorder. Using mouse models, they found that the low solid-cancer rate is attributable at least in part to an extra copy of a gene whose product blocks the formation of new blood vessels — a process called angiogenesis — that typically feed tumours (see page 1126). Ryeom tells *Nature* more.

How does this gene block angiogenesis?

We learned that DSCR1, a protein encoded by a gene on chromosome 21, suppresses one of the signals needed for new blood vessels to grow. When tumours are starting to form, they send out signals for blood-vessel growth. Humans can't usually fight that signal — it overrides the normal complement of DSCR1 that most of us have. But people with Down's syndrome have three copies of the gene, and we think that contributes to an ability to resist the tumour's signal.

Were you surprised by your findings?

Yes. When you do studies of cells in culture, you might express a gene 10,000-fold more than in its natural state to determine its function. But the mice we used had just one extra copy of the *DSCR1* gene and had more than 50% more DSCR1 protein than did the non-Down's-syndrome models. The question was whether that extra amount would be enough to shut down angiogenesis. It was.

What does this mean for human cancers?

We've learned from earlier clinical trials of anti-angiogenic agents that stopping the formation of new blood vessels once a cancer has taken root may come too late to arrest tumour development. We think that there is potential to develop an anti-angiogenic pill that could be taken every day — long before any cancer cells might begin to form — to stop tumours from growing in the first place. If we can show that DSCR1 blocks angiogenesis in the long term and has no side effects in mice, we will pass it on to industry. We're pursuing this now. ■

MAKING THE PAPER

Dewashish Upadhyay

Ancient rocks bear traces of Earth's tumultuous early years.

Cratons are the oldest, most stable parts of Earth's crust, and as such hold clues to Earth's early evolution. Dewashish Upadhyay, a geochemist now at the Indian Institute of Technology in Kharagpur, analysed the make-up of isotopes in rocks from India's Bastar craton and found that some of the rocks carry the signature of a differentiation event — the separation of materials with different geochemical properties. This event must have taken place during the first 400 million years of Earth's history, possibly when a magma 'ocean' covering the planet solidified.

The area encompassing the Bastar craton and the neighbouring Eastern Ghats Granulite Belt of southeastern India is rich in old rocks. The presence of alkaline igneous rock — a type of rock derived from Earth's mantle — marks the location where the Indian subcontinent split apart from another land mass 1.4 billion years ago and then collided with East Antarctica 1.1 billion years ago.

For his graduate project at the University of Bonn in Germany, Upadhyay made extended visits to the craton to collect samples and map the area. His initial analyses determined that the rocks were enriched in what he calls 'incompatible elements' — elements that prefer to remain in magma rather than become incorporated into crystallizing minerals.

During the course of this work, the geological field was stirred up by the discovery of rocks in southwest Greenland that provided evidence that a terrestrial magma ocean had crystallized to produce Earth's crust and mantle within 400 million years of the Solar System's formation. Isotopic analysis of the Greenland rocks had revealed an excess of neodymium-142 compared with terrestial standards, which indicated a separation of elements during that early, geochemically tumultuous time.

^{142}Nd is produced by the α-decay of samarium-146, a now-extinct nuclide. Major geochemical differentiation events, such as the formation and crystallization of a magma ocean, could have fractionated samarium and neodymium, producing reservoirs with contrasting ^{142}Nd compositions. Because the Greenland reservoirs were rich in ^{142}Nd, researchers started to look for a complementary reservoir with a ^{142}Nd deficit. And Upashyay wondered whether his rocks, rich as they were in incompatible elements, could have come from such a reservoir.

To test his suspicions, he needed to measure the ratio of ^{142}Nd to ^{144}Nd, a naturally occurring radioisotope that serves as a reference point. "If we can measure variation in ^{142}Nd, it implies there was some sort of a differentiation event that fractionated samarium from neodymium," Upadhyay says.

Upadhyay got the opportunity to do this analysis as a postdoc in Klaus Mezger's laboratory at the University of Münster in Germany. It revealed a deficit in ^{142}Nd in four out of seven rock samples. He and his coauthors concluded that the rocks were derived from a source formed during Hadean time — Earth's earliest geological aeon — and that some such Hadean reservoirs may be preserved in the mantle beneath cratons (see page 1118).

Now Upadhyay is back in India as a faculty member. He is working to set up his own lab, which will be the first in the country able to do the kind of isotopic analysis required to track elements — and so Earth's history. ■

FROM THE BLOGOSPHERE

Academic reports may sound alarm bells, but do not necessarily spur governments into action. In a post on Indigenus, *Nature India* editor Subhra Priyadarshini highlights two recent publications on the effects of climate change on vulnerable human populations (http://tinyurl.com/nqvq3t).

A report produced by the United Nations in collaboration with Columbia University in New York and the aid agency CARE International predicts large-scale migrations, and disruption to farming and water supplies, as the Himalayan alpine glaciers melt. Meanwhile, the London-based International Institute of Environment and Development, has published a book saying that a top priority, and no small task, is to "remedy deficits in infrastructure" in cities where "at least half of the population lacks piped water, sewers, drains, health care or emergency services," writes Pryadarshini.

She notes that past academic reports and press coverage of some of the world's first climate-change refugees have done little to change their plight: "The migrants from these sinking islands [in the Bay of Bengal] have not yet been recognised as vulnerable." ■

Visit Nautilus for regular news relevant to *Nature* authors ▸ http://blogs.nature.com/nautilus and see Peer-to-Peer for news for peer reviewers and about peer review ▸ http://blogs.nature.com/peer-to-peer.

nature

www.nature.com/nature

Vol 459 | Issue no. 7250 | 25 June 2009

Cheerleader or watchdog?

Science journalism is under threat. What can scientists do to help?

What should any researcher expect from a journalist beyond the keen intelligence needed to see the newsworthiness of the researcher's work, and the ability to spell his or her name correctly?

For some scientists, the answer is probably 'Not much'. Many tend to think of science journalism as a kind of public-relations service, existing purely to explain new scientific findings to the masses. They may well enjoy reading the results, and give points for a writer's ability to convey the excitement of a discovery, but they will mainly judge an article on its scientific accuracy.

On top of this, some will see science journalism as an ally, useful for shaping the public's understanding of science-related issues such as nuclear proliferation, stem cells or genetically modified crops — and, not incidentally, for making the case for a thriving research enterprise to public and politicians alike.

And a minority, moving beyond perceived self-interest, will point to the deeper value of journalism, which is to cast a fair but sceptical eye over everything in the public sphere — science included. This kind of scrutiny is easy for researchers to applaud when a news report questions dodgy statistics, say, or dubious claims about uncertainties in evolution. It is not so easy when the story takes a critical look at sloppy animal-research practices, overblown claims about climate change or scientists' conflicts of interest. But such examinations are to the benefit of the enterprise as a whole: society needs to see science scrutinized as well as regurgitated if it is to give science its trust, and journalists are an essential part of that process.

At the moment, unfortunately, journalism's future is far from clear. At the 6th World Conference of Science Journalists, which will be held next week in London, and of which *Nature* is a sponsor, there will probably be many attendees wondering if this is journalism's swan song. Readers — and small ads, once a reliable earner — are migrating to the Internet. New business models in which papers are given away have caused damaging dislocations in some markets, as in Denmark.

This contraction is perhaps particularly bad news for journalists with specialist beats such as science — the kind of journalists who need an informed understanding of what they are writing about, and know which experts can provide context, and where appropriate criticism, of new results. But publishers tend to see that kind of expertise as a luxury when money is tight, especially when the same space can be easily filled with material from press releases and wire services.

Scientists can do little to stem this bloodletting. But whatever they can do to engage with those caught up in it, and ensure that questioning and informed science journalism persists, will be worthwhile. If there is to be a transition to new — perhaps philanthropic — business models for in-depth reporting or new types of analytical media, science journalism will integrate into them all the better if scientists are taking an active interest in its health. And if the future of the media truly is a dire landscape of top-100 lists, shouting heads and minimal attention span, then such efforts might at least defer the grim end.

Even amid the turmoil, however, scientists can help ensure that reporting about science continues to be both informed and accurate. They can start by recognizing that they will increasingly be talking to reporters who have no experience in science at all. A good start would be to have a look at the advice for academics speaking to journalists provided by Brad DeLong and Susan Rasky at the University of California, Berkeley (http://tinyurl.com/nljleo). Scientific societies, meanwhile, should redouble their efforts to help journalists working under a tight deadline find the experts they need to understand a subject.

Looking farther down the road, the scientific community should work with journalism schools and professional societies to ensure that journalism programmes include some grounding in what science is, and how the process of experiment, review and publication actually works.

Science and journalism are not alien cultures, for all that they can sometimes seem that way. They are built on the same foundation — the belief that conclusions require evidence; that the evidence should be open to everyone; and that everything is subject to question. Both groups are comprised of professional sceptics. And whether it's directed towards an experiment or a breaking news story, each can appreciate the other's critical eye. ■

See also pages 1054, 1055, 1057 and 1061, and online at http://tinyurl.com/sciencejournalism and http://tinyurl.com/natureforum.

Raising the standards

A fledgling effort in China will show people what is happening on the environmental front.

China's central government has repeatedly declared its intention to clean up the environment, from the smoggy skies of Beijing to the scummy green waters of Lake Tai in the Yangtze delta. However, ensuring that intentions are translated promptly, fairly and efficiently into action across the provinces is often a problem.

Two non-governmental organizations (NGOs) may be about to change this. On 3 June, the Institute of Public and Environmental Affairs (IPE) in Beijing and the Natural Resources Defense Council in New York posted their first Pollution Information Transparency Index on the Internet. The index ranks 113 Chinese cities on how well they informed the public about pollution-related legal violations in 2008 — specifically, by measuring their adherence to the national environment-information disclosure measures that came

into effect on 1 May, 2008. The index incorporates data on how complete the cities' disclosures are, whether they are timely, what measures have been taken in response and whether the disclosures are user-friendly. The potential maximum score is 100. However, the average score for 2008 was only 30, and just four cities scored more than 60.

The two NGOs hope that the ranking will pressure local governments to recognize and be more responsive to the kind of situation that developed in Yangzong Lake in southern China. From 2001 to 2008, the lake — which supplied drinking water for 26,000 people — was polluted with arsenic by a chemical company, despite fines and promises of action from local officials.

The index is also intended to make role models out of the best-performing local governments. In each of eight categories, the IPE designated a top scorer. Taken together, they form a 'dream team' that scored 89.5 points. "This shows that a high level of information disclosure is possible in China," says Wang Jingjing at the IPE.

The system has limitations, however. It can only report on cases already recognized in official documents. That is not often an issue, says Wang, and an official source is usually available when the IPE hears of a complaint. But the group has no authority to push for further investigation if information is not available. Only 89 of 113 local environmental protection offices returned their calls.

A deeper problem is that the environmental laws themselves are not clear on crucial issues such as penalties for non-compliance. And in any case, the bar is pretty low. For example, Beijing ranks as one of the most transparent cities on the list, even though allegations that officials were manipulating data ahead of the Olympics have been met with denial, and little in terms of explanation (see *Nature* doi:10.1038/news.2009.578; 2009).

Although it is easy to be sceptical, local newspapers are paying attention to the index. In a face-saving society, local bureaucrats might now have to explain why they are at the bottom of the list. "The first thing to do is encourage public participation," says Wang. "But to do that people have to know what is happening."

The effort will also give central and local lawmakers a better idea of what is happening, and enable them to see whether laws need to be changed to ensure that China can make good on its intentions. ■

Orphan giant

Strong advocacy is needed if progress is to be made against tuberculosis.

It was difficult to avoid a sense of despair after last week's Pacific Health Summit, in Seattle, Washington. The meeting — an annual gathering of researchers, public-health policy-makers, drug regulators and heads of non-governmental organizations, industry and funding institutions — was focused this year on tuberculosis (TB). As presentation followed presentation, the overwhelming scale of the challenge became all too evident. The global economic downturn is exacerbating both the burden of the disease and the obstacles to finding resources to cope with it.

About one-third of the world's population carries the bacterium that causes TB, *Mycobacterium tuberculosis,* and roughly 10% of these people will go on to develop the disease. Because drug courses can last for a year or more, most people do not complete their treatment, and the growth of resistance is therefore inevitable. Strains of *M. tuberculosis* are now appearing that are resistant not just to the front-line drugs used in initial treatment but also to the second-line drugs used to treat people who have become resistant, as in the case of the 'extensively drug-resistant tuberculosis' first recognized in 2005. As a recent review article dauntingly but appropriately put it, responding to the issue of resistance requires "a comprehensive approach incorporating innovation from the political, social, economic and scientific realms" (M. Jassal and W. R. Bishai *Lancet Infect. Dis.* **9,** 19-30; 2009).

Happily, there are some signs of progress. The World Health Organization points out that existing drugs and diagnostics can still make a difference if applied properly. And major improvements in tackling the disease would come from a more joined-up approach with the diagnosis and treatment of HIV/AIDS, which often occurs together with TB.

Unhappily, many of these measures require trained staff, well-designed buildings and procedures that minimize the spread of infection — resources that simply aren't available in many regions of the world.

Despite donations of vaccines by industry, price-tiering in poorer regions, progress in clinical trials, advanced market-commitment mechanisms and increases in research funding by organizations such as the US National Institutes of Health, the scale of funding remains dwarfed by the challenges. What progress has been made in research simply clarifies just how inadequate our knowledge is. Researchers have only a basic understanding of how the bacterium affects the various parts of the body, and the heterogeneity of its make-up and behaviour is reflected in the spectrum of progressions from infection to active TB. Furthermore, little is known about how the human body responds to *M. tuberculosis* infection and the bacterium can mutate up to a thousand times faster as it adapts to antibiotic attack and other stressors.

"The timescale involved from research to the rolling out of new drugs and biomarkers is alarmingly large."

As Anthony Fauci, head of the US National Institute of Allergy and Infectious Diseases in Bethesda, Maryland, stated at the meeting in his overview of the science, generations of advances in research and technology have bypassed TB research. That is beginning to change, but the timescale involved from research to the rolling out of new drugs and biomarkers is alarmingly large given the immediacy of the threats and the complexity of the organism.

Why hasn't more progress been made? According to Margaret Chan, head of the World Health Organization, the field has been too isolated and inward-looking, and needs to learn lessons from the approach to HIV/AIDS by reaching out and finding highly effective champions. Easily said. But the nine million people who develop active TB every year could only agree that the need to capture the world's imagination and support is urgent. ■

RESEARCH HIGHLIGHTS

Birds of a feather

PLoS Biol. **7,** e1000132 (2009)

The size of flighted birds is limited by the demands of feather maintenance.

Sievert Rohwer at the University of Washington in Seattle and his colleagues studied 43 species of bird, assessing size, the length of flight feathers and timing of moulting cycles.

Flight-feather length is proportional, they say, to body mass raised to the one-third power, so that feather length roughly doubles with a tenfold increase in a bird's weight. But feather growth rate is proportional to body mass raised only to the one-sixth power. The trade-offs in time and energy required to replace long feathers therefore limit maximum bird size.

For a longer story on this research, see http://tinyurl.com/l3ge9b.

AFLO/NATUREPL.COM

ASTRONOMY

Honing the Hubble constant

Astrophys. J. **699,** 539–563 (2009)

Astronomers have measured the Universe's current rate of expansion to within 5% — twice as a precisely as before. Adam Riess of the Space Telescope Science Institute in Baltimore, Maryland, and his team used the Hubble Space Telescope to refine previous estimates of the Hubble constant, which relates the speed at which galaxies race apart to their distances from each other. They found the constant to be 74.2 kilometres per second per million parsecs.

Because true cosmic yardsticks are hard to find, astronomers chain together different observations to get the great distances needed to determine the constant. The researchers eliminated some sources of error using observations of 240 stars called cepheids, which provided precise distance measurements for seven galaxies. The new measurements are consistent with observations that the Universe is full of dark energy, a repulsive force that is accelerating the Universe's expansion — and changing the Hubble constant slightly with time.

NEUROBIOLOGY

Sweet memories

BMC Biol. **7,** 30 (2009)

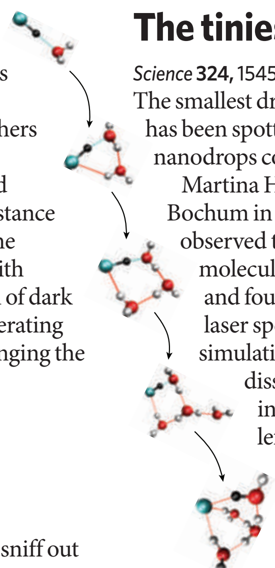
Honeybees have helped scientists to sniff out a trigger of long-term memory formation. Calcium has been linked to this type of memory for more than a decade, but only recently have researchers directly manipulated it in the honeybee (*Apis mellifera*) to show that it can activate the formation of such memories.

Jean-Christophe Sandoz and his colleagues at the University of Toulouse in France tested honeybees' memories by first training them to extend their probosces for a sugary reward when presented with a specific odour. They then exposed the bees to the same odour three days later to see whether the bees remembered it. Injecting the bees with a chemical that decreased calcium availability impaired the insects' ability to remember. Conversely, injecting chemicals that increased calcium levels improved their long-term memory.

CHEMISTRY

The tiniest acid drop

Science **324,** 1545–1548 (2009)



AAAS

The smallest drop of acid it is possible to make has been spotted within superfluid helium nanodrops cooled to below 1 kelvin.

Martina Havenith of Ruhr University Bochum in Germany and her colleagues observed the reaction between one molecule of hydrochloric acid and four of water using infrared laser spectroscopy. Using *ab initio* simulations, they found that the dissociation into H^+ and Cl^- occurs in a stepwise manner (pictured, left). A novel mechanism, which the authors term 'aggregation-induced dissociation', explains how this can happen even at such ultralow temperatures.

Such dissociations are key in chemistry, with relevance to processes ranging from reactions in polar clouds to spontaneous synthesis in space, the authors say.

NANOTECHNOLOGY

Mass spec goes mechanical

Nature Nanotechnol. doi:10.1038/nnano.2009.152 (2009)

Mass spectrometry analyses molecules on the basis of the mass-to-charge ratios of their ionic fragments. The method can be difficult to use for substances that are hard to ionize or only available in minute amounts. Michael Roukes and his colleagues at the California Institute of Technology in Pasadena now demonstrate a nanoelectromechanical device that can directly measure the mass of individual molecules.

Inside the device, a tiny suspended beam vibrates at high frequency and acts as a sensor.The authors blasted the device first with gold nanoparticles, and later with bovine serum albumin.

When molecules stick to the sensor's surface, they drastically lower its vibrational frequency. By measuring the change, the mass of individual molecules can be established, the authors say, and this could potentially be done without the use of ionization.

ATMOSPHERIC SCIENCE

Rain on physics

Geophys. Res. Lett. doi:10.1029/2008GL037111 (2009)

By observing approximately 64,000 raindrops, researchers have determined that some can fall by as much as an order of magnitude faster than their terminal velocity — the limit at which the downwards force of gravity equals the upwards force of drag.

Fernando García-García at the National Autonomous University of Mexico in Mexico City and his colleagues show that when large

drops break up, the fragments continue to move at the faster speed of the larger parent drop for some time before slowing down to their own terminal speeds.

The results could improve the representation of rain physics in weather models.

BIOLOGY

A light touch

Science **324,** 1580–1582 (2009)

Merkel cells were discovered in the late nineteenth century and have long been suspected of playing a part in sensing light touch. However, this has never been entirely settled experimentally.

Ellen Lumpkin and Huda Zoghbi at Baylor College of Medicine in Houston, Texas, and their colleagues selectively deleted a gene encoding a key transcription factor, ATOH1, in cells from the body skin and foot pads of mice, where Merkel cells are normally found. Merkel cells were absent from these areas, and by testing the response of different nerve types the researchers found a "conspicuous loss" of a specific type of touch response.

Further work will be needed to determine whether the cells are sensory receptors or have some other role in touch.

ORGANIC CHEMISTRY

Give and take

Nature Chem. doi:10.1038/nchem.248 (2009)

Rare compounds in which carbon — the archetypal organic element — seems to behave like a metal may be more widespread than previously thought.

In typical organic compounds, a carbon atom forms four covalent bonds by sharing its four most accessible electrons. But Alois Fürstner and his colleagues at Germany's Max Planck Institute for Coal Research in Mülheim an der Ruhr argue that in some newly synthesized and previously known compounds, carbon's bonding behaviour is best described as donating pairs of electrons to, or accepting pairs of electrons from, neighbouring atoms — behaviour seen in positively charged metal ions.

This peculiar situation, which blurs the traditional divide between organic and inorganic chemistry, could form the basis of new types of catalyst.

L. CARDINALI

GENETICS

Depression link revoked

J. Am. Med. Assoc. **301,** 2462–2471 (2009)

An analysis of the literature affirms doubts about a much-hailed gene–environment link for depression.

In 2003, researchers found that individuals with specific versions of the serotonin receptor gene *5-HTTLPR* are more susceptible to depression when challenged by stressful life events (*Science* **301,** 386–389; 2003). The finding made intuitive sense, and many studies attempted to replicate and build on the results.

Now, Kathleen Ries Merikangas at the National Institute of Mental Health in Bethesda, Maryland, and her colleagues have evaluated the original study together with 13 others that closely replicated its conditions, re-coding data where necessary to match the original. Looking at a total of more than 14,000 study participants, they find no association between *5-HTTLPR* type and risk of depression, regardless of life events.

COGNITIVE SCIENCE

The tool becomes him

Curr. Biol. **19,** 10.1016/j.cub.2009.05.009 (2009)

Tool use can alter people's perception of the size and position of their body parts.

INSERM's Alessandro Farnè and Lucilla Cardinali at Claude Bernard University in Lyon, France, and their colleagues had volunteers use a mechanical grabbing arm (pictured left) that extends reach by 40 centimetres. Measuring arm movements with a high-resolution motion-tracking system, they showed that the volunteers reduced the velocity and acceleration of movements to grab or point at an object after a period of using the tool.

When blindfolded, the volunteers also estimated their arms to be slightly longer after tool use, confirming a long-held hypothesis that body schema — the sense of where one's body parts are in space — is plastic.

For a longer story on this research, see http://tinyurl.com/mwjpow.

JOURNAL CLUB

Michael S. Fuhrer
University of Maryland, College Park

A physicist peels back the layers of excitement about graphene.

Graphene is an atom-thick sheet of carbon in which electrons behave as if they have no mass. Atomic carbon layers have been grown epitaxially — that is, perfectly aligned with atoms in an underlying crystal surface — on metals and semiconductors for decades, so why the fuss lately?

Well, in the past few years much work in this field has revolved around graphene obtained by 'exfoliating' or peeling it from graphite. By mounting exfoliated graphene on insulating silicon dioxide, researchers observed a half-integer quantum Hall effect, an anomalous measurement that stems from the existence of a Landau level — the quantized orbit of electrons in a magnetic field — at exactly zero energy, a signature property of massless electrons.

But exfoliated graphene is dirty, lumpy and tiny (the biggest pieces are still a tenth of a millimetre in diameter). I wondered whether the older technique of epitaxial growth could produce a better material. In May, a group led by Joseph Stroscio showed that it could (D. L. Miller *et al. Science* **324,** 924–927; 2009). Using scanning tunnelling spectroscopy to study epitaxial graphene on the surface of silicon carbide in a magnetic field, they showed that epitaxial graphene is extraordinarily clean and flat, and clearly exhibits the zero-energy Landau level. For the first time in any material, epitaxial graphene allows direct observation with atomic resolution of the behaviour of electrons in quantized Landau levels, opening a new window on the quantum Hall effect.

Researchers using other techniques, such as cyclotron resonance and photoemission, are also reporting an astonishingly clean electronic system in epitaxial graphene. Experiments on conductivity remain a challenge, but epitaxial graphene seems to have a bright future.

Discuss this paper at **http://blogs.nature.com/nature/journalclub**

NEWS

US human spacefaring questioned

Review panel takes a hard look at NASA's goal of returning astronauts to the Moon.

When it comes to how NASA should spend its money putting people in space, just about everyone has an opinion — some of them strident. But at the moment, the opinions that matter most are those in a ten-person committee that listened quietly as key players pleaded their case at a meeting on 17 June in Washington DC.

Led by retired Lockheed Martin chief executive Norman Augustine, the committee has been tasked by President Barack Obama to review the United States' plans for humans in space. It includes a mix of aerospace executives, astronauts, engineers, a retired general and two scientists: Christopher Chyba, a planetary scientist at Princeton University in New Jersey, and Earth scientist Charles Kennel, chair of the National Academies' Space Studies Board.

The meeting — its first public hearing — showcased the dire budget constraints the agency faces as it struggles to replace the space shuttle, which will be retired by 2010, with rockets that could return astronauts to the Moon. The hearing also made plain the number of people who feel that NASA's existing designs aren't the best way to do it.

P. E. ALERS/NASA/GETTY IMAGES

Panel chair Norman Augustine.

Several groups presented alternatives to the committee and said they could be completed faster and more cheaply than the Ares rockets in NASA's planned Moon programme, called Constellation (see graphic). Michael Gass, chief executive of United Launch Alliance in Littleton, Colorado, told the panel that his company's Delta IV cargo rocket could be modified to carry a crew capsule. Stephen Metschan, representing a group that claims to include NASA engineers working in their spare time, presented designs for a system called Direct, based on existing shuttle hardware.

Steve Cook, Ares project manager at Marshall Space Flight Center in Huntsville, Alabama, says such proposals may reflect lingering bitterness over lost contracts. In 2005, a NASA study selected what became the Constellation system as the best method for getting astronauts past low-Earth orbit and back to the Moon. "When decisions are made, somebody wins and somebody loses," Cook says. "Those that didn't win don't like that."

Budget stretch

Getting beyond low-Earth orbit may be too much of a stretch for NASA's budget, however, regardless of which rocket is used to do so. Cook says that the agency has already spent US$10 billion of the $35 billion needed to complete the *Orion* crew capsule and the Ares I rocket, which together could fly astronauts to

New protein structures replace the old

Protein structures are getting regular makeovers with the help of 're-refinement' software developed by Dutch structural biologists.

The Protein Data Bank (PDB) holds nearly 53,000 three-dimensional structures of protein molecules and nucleic acids that have mainly been deciphered through X-ray crystallography. Most journals, including *Nature*, require such data to be deposited in the PDB if a paper with a protein structure is to be published.

But some structures are not as accurate as they could be. The data bank began in 1971, and the ability to analyse crystallographic data has improved dramatically since then.

"There are definitely errors in the PDB," says crystallographer Nenad Ban at the Swiss Federal Institute of Technology Zurich.

This has consequences for scientists who want to use the PDB to look for sites on proteins to target with small-molecule drugs, or to feed the data into molecular dynamics simulations. More profoundly, a wrong structure in the data bank could also trigger wrong ideas about how the protein works.

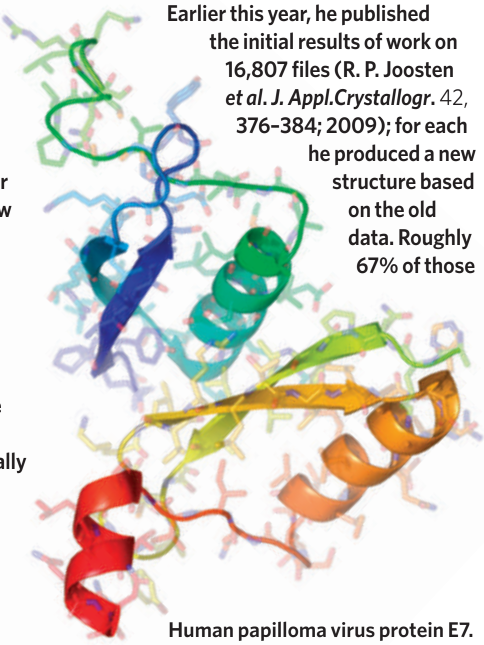
To help, Gert Vriend at Radboud University Medical Centre in Nijmegen, the Netherlands, and his colleagues are writing software that they hope will eventually automatically re-refine, at the click of a mouse, all the data deposited in the PDB.

So far, Vriend has gone through 38,000 data files with his PDB_REDO program. Earlier this year, he published the initial results of work on 16,807 files (R. P. Joosten *et al. J. Appl.Crystallogr.* 42, 376–384; 2009); for each he produced a new structure based on the old data. Roughly 67% of those new structures were better than the original structure as measured by a quantity known as *R*-free, used by crystallographers to determine structure quality.

The automated program could also decipher other problems caused by human error: data deposited with the wrong labels, for example, in which intensity of a signal is labelled as amplitude. "If the intensity and amplitude are swapped, the structure doesn't make sense," says collaborator Robbie Joosten.

Vriend runs his program on all new entries in the PDB every two weeks and sends the PDB a monthly report flagging problems. Administrators can correct small problems like names of labels being swapped; bigger problems are added to an ongoing list of things to fix.

Maintenance for the PDB

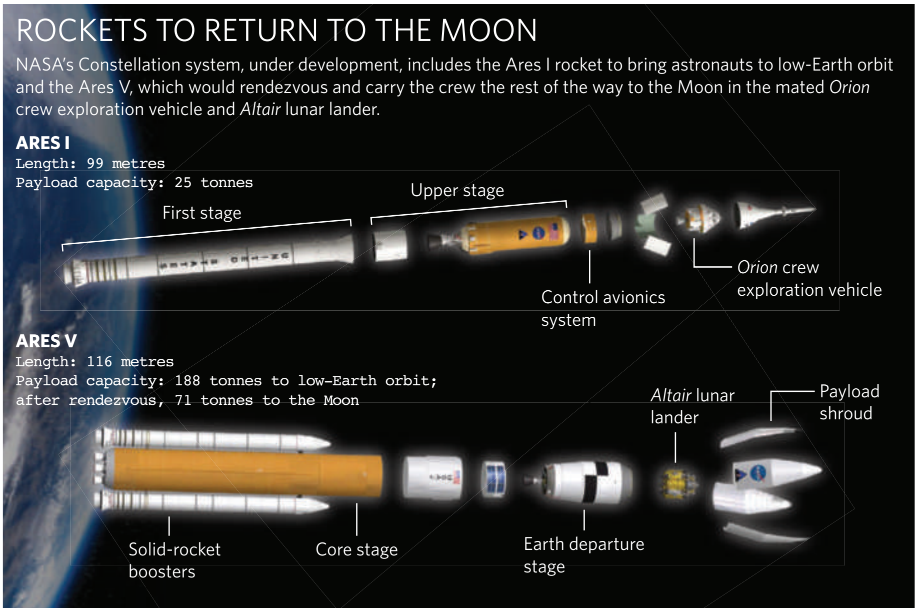


LAGUNA DESIGN/SPL

Human papilloma virus protein E7.

APOLLOPLUS40 ON TWITTER
Nature News relives the Apollo 11 Moon mission.
http://twitter.com/apolloplus40

NASA/JPL/USGS



ROCKETS TO RETURN TO THE MOON

NASA's Constellation system, under development, includes the Ares I rocket to bring astronauts to low-Earth orbit and the Ares V, which would rendezvous and carry the crew the rest of the way to the Moon in the mated *Orion* crew exploration vehicle and *Altair* lunar lander.

the International Space Station, by 2015. He estimates it will cost $100 billion to complete the Ares V, which would pick up astronauts in orbit and carry them to the Moon, by 2020.

The committee will examine not only the means of spaceflight but also its ends. Augustine says his team will "look at the full spectrum of possible destinations" other than the Moon. It has also been asked to evaluate using the space station after 2015.

Augustine says he has received lots of correspondence from scientist friends and acknowledges that many of them feel that putting astronauts on the space station and the Moon isn't all that useful scientifically. But others disagree, he says. Advocacy groups such as the Planetary Society in Pasadena, California, have lobbied for other destinations, such as near-Earth asteroids, as intermediate steps to the human exploration of Mars. Furthermore, some astronomers believe that astronauts — and the beefier rockets that are needed to carry them beyond low-Earth orbit — will be instrumental in launching large space telescopes, or building and repairing them in space.

Florida Senator Bill Nelson (Democrat), one of NASA's powerful patrons in Congress, told the committee that Obama would probably heed their advice. "You come to the table with extraordinary influence," he says. Nelson, an advocate for increased NASA funding, exhorted them to find the most meaningful programme goals, even if that meant ignoring the budget constraints given to them.

Although the committee members spent most of the day listening rather than talking, Augustine says his group will not be bashful in its report if it finds that NASA can't do much worthwhile within its budget. "Just because something's cheap doesn't mean it's worth doing," he says.

The committee's report is due in August. Former astronaut Charles Bolden, who has been nominated as the next NASA chief, would implement any changes that result from it. ■

Eric Hand

---

is spread over three sites: the Research Collaboratory for Structural Bioinformatics based at Rutgers University in Piscataway, New Jersey; the European Bioinformatics Institute (EBI), in Hinxton, UK; and the Japan Science and Technology Agency in Tokyo.

Helen Berman of Rutgers, who runs the US part of the PDB, says that the data bank welcomes efforts to improve deposited data. "This is exactly the vision we had when we started," she says. Staff at the data bank run a standard set of quality checks before depositing a structure and flag any problems to the researcher who submitted it, but "we're not the data police". The data bank doesn't reject an entry even if its advice on improving a structure is ignored.

Vriend occasionally contacts the scientists who deposited data that he has refined, with mixed reactions. "Sometimes people are very grateful, and sometimes they are insulted," he says.

"Sometimes people are very grateful, and sometimes they are insulted."

Occasionally, the program can cause researchers to change interpretations of their data. Annalisa Pastore, a molecular biologist at the National Institute for Medical Research, London, UK, recalls asking Vriend to validate a protein structure she had worked out from her nuclear magnetic resonance (NMR) data. Vriend told her she had got it wrong, and she took a closer look. It turned out that she wasn't wrong, but that she had uncovered a histidine residue that was unusually buried within the protein. "Gert correctly focused our attention to this residue," she says. "In the end we could definitely say the structure was right."

Pastore says researchers might be more careful about submitting their structures to the data bank if they think re-refinement software might be checking up on it.

Vriend is not the only one looking at data-bank quality. At the Lawrence Berkeley National Laboratory in California, computational biologist Paul Adams is testing the PHENIX crystallography software he develops on PDB data before sharing the software with other academics or licensing it to companies. "We [want to] make sure the software we are giving to people can do the right thing," he says. Adams doesn't make his results public, but says he has noticed that his software often improves an original structure assignment.

Vriend hopes his re-refinement software will eventually be linked to the PDB so that a user could click through from the data bank to obtain the most up-to-date protein structure. Gerard Kleywegt, who took over the European PDB operations at the EBI latelast month, says that this will probably be implemented at some point.

Even so, the software is not sophisticated enough to automatically fix problems that are more than cosmetic, Kleywegt says. More serious problems, such as amino-acid side chains that have been assigned to the wrong location, require manual intervention.

"I see this as a first step," he says. ■

Katharine Sanderson

CDC

Pandemic virus update

ASIA: Australia and the Philippines both report deaths of people infected with the novel H1N1 virus — the first in the Asia-Pacific region.
IRAN: reports its first swine-flu case, a 16-year-old boy visiting from the United States.

Climate burden of refrigerants rockets

Environmentalists push for tougher regulation of chemicals meant to help the ozone layer.

Modern refrigerants designed to protect the ozone layer are poised to become a major contributor to global warming because of their future explosive growth in the developing world, scientists report this week.

Hydrofluorocarbon chemicals (HFCs) were developed to phase out ozone-depleting gases, in response to the Montreal Protocol. But they can be hundreds or thousands of times more powerful than carbon dioxide as greenhouse gases in trapping heat. HFCs are deployed in refrigerators and air-conditioning units, and their use is poised to grow in the coming decades.

In the new study, a team led by Guus Velders at the Netherlands Environmental Assessment Agency in Bilthoven analysed the latest industry trends and then modelled HFC production to 2050. Their results suggest that HFC emissions could be the equivalent of between 5.5 billion and 8.8 billion tonnes of carbon dioxide annually by 2010 — roughly 19% of the projected CO_2 emissions if greenhouse gases continue to rise unchecked (G. J. M. Velders *et al. Proc. Natl Acad. Sci. USA* doi:10.1073.pnas.0902817106; 2009).

The new numbers will fuel the efforts of environmentalists and others who have been pushing for aggressive new HFC regulations. Manufacturers could shift towards using HFCs with the lowest climate impact during the transition to a new generation of refrigerants — still under development — that affect neither the ozone layer nor the climate.

"Now is the moment to make a decision to steer this in a direction that you want."

"Now is the moment to make a decision to steer this in a direction that you want," Velders says. "The developing world is already in the transition to HFCs."

Although it makes no policy recommendations, the study could play into an ongoing political debate on regulating the chemicals. HFCs currently fall under the umbrella of the Kyoto Protocol on climate change, but advocates say the fastest and cheapest way to handle them is under the ozone treaty. Montreal delegates plan to discuss the issue when they meet in Geneva next month.


Air-conditioning in cars is a major HFC source.

Durwood Zaelke, who heads the advocacy group Institute for Governance and Sustainable Development in Washington DC, says the Velders study confirms the potential benefits of regulating HFCs under the Montreal agreement. It will, he says, help build momentum as the delegates move towards a decision in November.

Deficit dooms Swedish gene institute

Cost overruns have sliced into the historically important biology programme of Sweden's Lund University, culminating in a departmental reorganization that has some scientists fuming.

The Institute of Genetics, where in the winter of 1955-56 the correct number of human chromosomes was established (J. H. Tjio & A. Levan *Hereditas* 42, 1–6; 1956), is being disbanded and four researchers laid off in a bitter reshuffle.

"Some researchers feel they can't afford to be nice to their colleagues," says Sven-Axel Bengtson, who directs the university's 274-year-old Museum of Zoology. "When food is scarce, the horses will bite."

Úlfur Árnason will lose his office.

Acrimony peaked last week, when an emeritus professor was notified that he would lose his office by 30 June. Úlfur Árnason, who has been associated with the university for more than 50 years, was told he was to be ejected shortly after he sent a lengthy critique on the biology reorganization to the university's top executive, rector Per Eriksson. Árnason was among the few who openly questioned the proposal, contending it would undermine science at the university, reduce enrolment and cause revenue to drop further. He also noted that the vision group directed by zoologist Dan-Eric Nilsson, ranked 'outstanding' in a 2003 external review, was being cut, whereas a plant-research group criticized in the same review as being without focus was left largely untouched. Susanne Widell, who leads the plant group, is directing the biology reorganization; she did not respond to an interview request.

Árnason became an emeritus professor four years ago, as Swedish researchers do at age 67. Typically, such professors are permitted a university office if they continue research, as Árnason does. He is fighting the ousting.

HAVE YOUR SAY
Comment on any of our News stories, online.
www.nature.com/news

ALAN SCHEIN PHOTOGR./CORBIS

The Montreal treaty, which came into force in 1989, has implementation experts in virtually every country and has already succeeded in reducing 96 chemicals by 97% each, Zaelke says. "It's a winning record, and we need to give the treaty this shot."

In 2007, Velders' team looked at the effect of the Montreal Protocol and found that its incidental greenhouse-gas reductions — equivalent to 11 billion tonnes of CO_2 annually by 2010 — are five to six times greater than those of the Kyoto Protocol. Montreal participants had already agreed to phase out chlorofluorocarbons by 2010 and hydrochlorofluorocarbons by 2040, but the Velders study helped convince them to accelerate phasing out the latter by a decade in order to capture the climate benefits.

The latest study carries that work forward and suggests that the problem posed by HFCs could be several times larger than projected by the Intergovernmental Panel on Climate Change (IPCC) in its 2007 assessment. Using the IPCC's baseline economic scenarios, the work assumes that refrigerant technologies will be deployed at levels roughly equal to those in the developed world today. By mid-century, emissions in currently developing countries would rise to levels eight times higher than those in developed nations.

Venkatachalam Ramaswamy, a coordinating lead author of the IPCC chapter covering HFCs who was not involved in the study, called the paper "a very good piece of work".

"It elevates the importance of HFCs in terms of climate forcing to a level higher than we may have thought initially," says Ramaswamy, a scientist at the National Oceanic and Atmospheric Adminstration's Geophysical Fluid Dynamics Laboratory in Princeton, New Jersey.

Still, he says, it "doesn't significantly detract from the attention that CO_2 and methane should be getting." ■

Jeff Tollefson

"This is very extraordinary, to be expelled from the faculty," he says.

"It is quite astonishing," agrees Lennart Ståhle, a senior adviser at the Swedish National Agency for Higher Education in Stockholm.

At its peak, the genetics institute was home to about 80 researchers. The two dozen researchers there now will be reposted elsewhere within the university, and the existing genetics labs largely turned into offices.

University officials and researchers say that the reorganization was brought on by a series of compounding problems, many of which other universities also face, Ståhle says. At Lund, officials say, too many young researchers were kept on without outside funding to support them. The annual budget deficit of the Department of Cell and Organism Biology, which oversees the genetics institute, has been ballooning over the past couple of years, increasing by up to 100,000 Swedish kronor (US$12,500) a week, and is expected to hit 19 million kronor by year-end.

The science faculty board has approved a 10-million-kronor subsidy , which, combined with staff reductions, is intended to eliminate the deficit in three years.

"This is very extraordinary, to be expelled from the faculty."

Ecologist Torbjörn von Schantz, dean of the science faculty, said he "deeply regretted" the "tragic" cuts in staff, which he acknowledged will reduce future graduate-student admissions. But he added that the steps are necessary to save money.

"We see the restructuring of biology in Lund as a key to addressing new areas of research," he says.

However, several other professors privately questioned whether students would want to come to a department in turmoil. ■

Rex Dalton

G8 cancels science parley

The Italian government has caught scientists off guard by cancelling the G8 science and technology meeting that was to have begun on 25 June.

The meeting would have brought together science ministers of the G8 nations to discuss issues such as how to monitor climate change and maintain environment-friendly energy supplies.

It was one of ten ministerial meetings to prepare for the main G8 summit, in L'Aquila, Italy, in July. The summit is now left without a formal mechanism for introducing science input.

"We were surprised and disappointed by the cancellation," says James Wilsdon, director of science policy for Britain's Royal Society. The society is one of 13 national academies — one from each of the G8 countries, plus one each from South Africa, India, Mexico, China and Brazil — that had prepared a joint statement on climate and energy policy to funnel through the meeting.

The Italian government gave no clear reason for the cancellation, the first time a planned ministerial meeting has been cancelled in advance of a G8 summit. Italian newspapers have reported that government officials are referring variously to concerns about security and failure to complete translation of relevant preparatory documents .

Giuseppe Fioroni, science spokesman for Italy's opposition Democratic Party, says that the cancellation — which he puts down to infighting between ministries in the Italian government over the core themes of the research G8 meeting — is "humiliating for Italy as the host country".

The Italian Ministry of Education, Universities and Research did not respond to interview requests.

An official at the German research ministry said they had been told that the meeting would be reconvened in some form this autumn. "It is always good when G8 ministers meet," he said. "It does not have to be before a summit."

But Fioroni says a research meeting has meaning only if it takes place before the main summit. ■

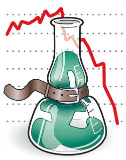
Alison Abbott

Florida universities hit by funding crisis

Stimulus money provides lifeline for cash-strapped public colleges.

This March, Michael Perfit, chair of the geology department at the University of Florida in Gainesville, learned that budget cuts might slash 11 of the 24 faculty and staff positions in his group. "It was like someone saying your whole family was killed in a car accident," he says. "I thought, 'This will destroy us.'"

Perfit's department ultimately dodged the full car wreck, thanks to a combination of factors that blunted — at least temporarily — the planned massive reductions in state funding for university research. But happy conclusions are not the norm this year. Like other US sunbelt states, including Arizona (see *Nature* **458,** 14–15; 2009) and California (see *Nature* doi:10.1038/news.2009.553; 2009), Florida has learned the hard way that real estate and tourism are not stable economic anchors in a recession.



Florida's 11 public universities have already endured years of severe budget cuts (see 'Marine biologists say save our subs'), and the fiscal year that begins on 1 July will be worse. Researchers embarked this spring on last-minute lobbying to save their departments. They are also looking to new funding streams — particularly the $787-billion economic stimulus bill passed by Congress in February. The new money has prevented some cuts, and others hope that stimulus grants will help researchers directly.

University leaders say they are already cutting into the meat of research budgets. The University of Central Florida in Orlando announced on 4 June that it plans to phase out five departments, including statistics and engineering technology; administrators say the decision was made on factors such as teaching and research quality, student demand and costs. Even that was just one $6-million component of a plan to deal with a $38-million cut in recurring funding.

At the University of Florida, the College of Liberal Arts and Sciences — which contains Perfit's department — initially feared it would lose 10% of its $93-million annual budget. In February, departments began looking at how to plan for such a worst-case scenario. The dean of the college, Paul D'Anieri, recommended the geology department, the operating budget of which was already down 65% since 2005, as one group to look at for potentially deep staff cuts.

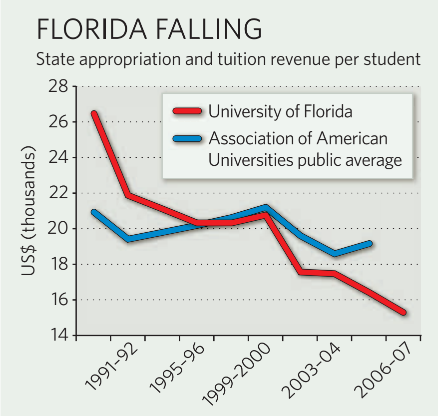
David Richardson, the university's associate dean of research, says when deciding what to cut, the administration took into consideration a range of factors such as size, number of students served and productivity. But decisions were ultimately constrained by the fact that tenured university faculty belong to a union that negotiates contracts collectively, which meant that non-tenured positions were easier to eliminate. "At some level, frankly, those lay-off orders were determined more by that than by choices made on the basis of what is best," says Richardson.

By March, Perfit was looking at losing all of the ten non-tenured faculty, plus one tenured position in his team. He sprang into action, soliciting friends of the programme in fields ranging from oil exploration to the environmental industry to write 75 letters of support to administrators. Without support staff to run the department's mass spectrometers, he argued, grant-funded projects in geology and other departments would be jeopardized. And if junior faculty left, their grant money might leave with them. "For me it was more than just the people; it was like my child was being destroyed," says Perfit. "I was fighting with all I had."

The impact of the lobbying campaign is difficult to gauge. But, says Richardson, "those efforts clearly focused attention on the role of these academic and research programmes."

A number of factors worked in Perfit's favour, primarily that, in the end, the college of arts and sciences suffered only a 3.8% cut to its budget instead of 10%. The administration did not have to resort to lay-offs in the department, but Perfit lost two positionswhen faculty who were already leaving were not replaced.

An influx of stimulus funding played a major role in the outcome, reducing the required cut in the total university budget from $42 million to $32 million. In addition, a bill passed on 1 June allows Florida's public universities to raise tuition fees by up to 15% per year if they choose — an option that had been barred by law until now. However, barring some source of replacement funding, deeper cuts will still have to be made in the next budget cycle.



UF NEWS BUREAU

Marine biologists say save our subs

Two industrious submersibles that have logged more than 8,500 dives between them could undergo enforced retirement as the management of the Harbor Branch Oceanographic Institute (HBOI) in Fort Pierce, Florida, considers selling their mother ship.

Many marine scientists are agitated at the thought of losing access to the Johnson Sea Links (JSLs), which have been a part of the HBOI since its founding in the 1970s. More than 2,000 scientists and non-scientists worldwide have signed a petition asking the state to support the JSLs.

In the face of state-wide budget cuts (see main story), Florida Atlantic University (FAU) at Boca Raton, which administers the HBOI, is reviewing several scientific departments, including the institute. In 2004, funding from the estate of its founder, Seward Johnson, was cut off; since then it has floundered financially, and in 2007 it merged with FAU. Harbor Branch still struggles to fund its US$20-million annual budget.

In May, HBOI and three other institutions were awarded a joint $22.5-million grant over five years from the National Oceanographic and Atmospheric Administration (NOAA) to establish a cooperative institute for ocean exploration, headquartered at the HBOI. But because only a portion of the grant will come to the HBOI, it is not sufficient to fund sub operations, says Peter Tatro, the institute's associate executive director. NOAA spokeswoman Jana Goldman says that specific tools such as the JSLs are not required by the grant, as long as the research goals are carried out.

Because of the budget woes, one JSL has been mothballed since 2006, and the second does not get

R. CARSON/UF NEWS BUREAU

The University of Florida applied for $75 million in federal money to counteract flagging state support.

Stimulus funding will become even more important to universities in the immediate future, as agencies such as the National Science Foundation begin awarding grants to universities now applying directly for stimulus research funding. The University of Florida has submitted some 300 such proposals in the past three months for $75 million in research and infrastructure grants.

"Every major research university is doing the same thing," says Win Phillips, the university's vice-president of research. "It's an overwhelming opportunity." The smaller University of South Florida has churned out 132 proposals for nearly $100 million in grants.

With so many applications submitted, success rates are likely to be even lower than usual, and some are questioning whether all the paperwork will pay off (see *Nature* **459,** 763; 2009). Still, says Richardson, "the grants that do come will be welcome, and will allow us to keep research groups going to a larger extent than we could have without them."

Ultimately, he hopes the benefits will be more profound aong scientists struggling for any ray of hope in the current economic climate. "I think that there is going to be a sense in the academic profession that government support for research is on the upswing," he says, "and that will encourage people to remain dedicated to their research goals."

Mark Schrope

enough commissions to earn its keep. Only one ship, the RV *Seward Johnson*, has the crane needed to launch a JSL, and to keep both ship and one sub accessible costs $3 million a year. To break even, the *Seward Johnson* — which can also carry out non-JSL missions — must spend 250 days at sea each year; in 2007, that number was 160, says FAU provost John Pritchett.

Concern over the subs' fate led some HBOI scientists to organize the online petition. "We didn't feel that enough was being done to keep those resources," says biologist Chip Baumberger. He hopes the petition will convince Florida legislators, or another funding agency, to support the subs.

The JSLs' useful features include an acrylic bubble giving a panoramic view, enough room for four people and hundreds of kilograms of equipment, and the manoeuvrability to chase down and collect live specimens. "All of the other submersibles have some capabilities, but the combination of all that JSL has, it really is unrivalled," Baumberger says. The JSLs salvaged the booster rocket with the faulty seal from the space shuttle *Challenger*, and have discovered several coral reefs and more than a hundred new species.

Some scientists are optimistic about future funding from government agencies or from the federal stimulus package, and predict the subs will again do good business. "It would be very shortsighted to sell the ship," says John Reed, a coral-reef biologist at the HBOI.

Other options include leasing the *Seward Johnson* or modifying the JSLs so they can launch from any ship. No decision has been made, Tatro says; the review will be completed this autumn.

"I'd like to find a federal agency that says, 'We'd like to have a manned submersible programme, and we'll pay for it.'" **Amber Dance**

NATURE|Vol 459|25 June 2009

Biology society narrows Chinese rifts

TAIPEI

Warming scientific relations between Taiwan and mainland China were on view last week at a meeting in Taipei of the Society of Chinese Bioscientists in America (SCBA).

Historically a Taiwanese-dominated academic networking group, the society last week announced its first president-elect born and raised in mainland China: Xiao-Fan Wang, a cancer biologist at Duke University in Durham, North Carolina.

And a bill introduced in the Taiwanese legislature in January could mean that many more mainland Chinese scientists will be soon be living and working in Taiwan. The legislation would permit Taiwanese universities to recognize university degrees from mainland institutions such as Peking University and Tsinghua University. "It's ridiculous that we don't," says Ji-Wang Chern, dean of research and development at the National Taiwan University in Taipei.

Carmay Lim, a computational biophysicist at the Academia Sinica in Taipei, says she has not been able to offer PhD positions to mainland students whom she wanted to stay on in her lab. "Because of the politics between Taiwan and [mainland] China, scientists like myself lose good students to universities in other parts of the world," she says.

Few details of the bill are publicly available, but Chern says it is likely to be adopted. The proposal aligns with generally better relations between mainland China and Taiwan, especially since Ma Ying-jeou became president of Taiwan in May 2008.

The easing of tensions could be seen in the relaxed atmosphere of the SCBA meeting. Society scientists recall past episodes in which a senior education ministry official from Beijing stormed out after not being allowed to move her talk ahead of a Taiwanese scientist, and a 2001 meeting where Nobel laureate Yuan Tseh Lee suggested that China was hurting chances for diplomacy and scientific collaboration by building up missiles aimed at Taiwan. His counterpart from the Chinese Academy of Sciences defended China's actions as a valid protection against imperialism. "There has been nothing like that here," says Chung S. Yang, a cancer researcher at Rutgers University in Piscataway, New Jersey, and former SCBA president.

New society president Kuan-Teh Jeang, a virologist at the US National Institutes of Health in Bethesda, Maryland, says the SCBA will be changing dramatically in ways that could help to unite Chinese scientists. He hopes to turn it into an organization with more lobbying power, an enhanced role in advising government and academic institutions on interactions with Chinese institutions overseas and, possibly within a year, its own online open-access journal.

"We need to be ready to extend beyond niche special interests," he says. "We should participate in the mainstream." ■

David Cyranoski

Chemistry publisher in move from print journals

The American Chemical Society (ACS) is taking steps to turn most of its academic journals into online-only publications.

According to a letter seen by *Nature* — sent to ACS associate editors by Susan King, senior vice-president of ACS Publications in Washington DC — the move has been prompted by the "accelerated decline in demand for print subscriptions and the diminishing financial return from the print format". King foresees a "move beyond print to an electronic-only scientific publishing environment".

To save money, most ACS journals will, from July, begin printing two pages of reduced text sideways on each page. At the same time, subscribers will be offered incentives to switch to online-only access. In 2010, ACS members will no longer be able to buy print subscriptions of journals.

For a longer version of this story, see http://tinyurl.com/llae53

More financial woes for climate satellites

NOAA

Ineffective management and cost overruns are plaguing an already troubled satellite system for tracking weather and climate, the US Government Accountability Office (GAO) reported to Congress on 17 June.

The National Polar-orbiting Operational Environmental Satellite System (NPOESS; illustrated right) was conceived in 1994 as a three-way project between the Department of Defense, the National Oceanic and Atmospheric Administration (NOAA) and NASA. "This tri-agency experiment has been an epic failure," says the GAO's David Powner.

In 2002, NPOESS was supposed to cost $7 billion for six satellites, with the first to launch in 2009 (*Nature* **450,** 782–785; 2007). Auditors now think that the current estimate — $14 billion for four satellites beginning in 2014 — is at least $1 billion off the mark. A separate report, conducted by an independent review team for NPOESS management, recommends that the White House appoint a top-level decision-maker, and suggests that NOAA should be made lead agency on the system.

Researchers urge action on medical-isotope shortage

US and Canadian medical researchers last week intensified their calls to tackle a worldwide shortage of radioisotopes for medical imaging.

Sudden dips in isotope supply occurred regularly in 2007 and 2008 as elderly reactors were shut down for repairs. The issue flared up again this year after an unscheduled 14 May shutdown of a large 52-year-old nuclear-isotope-production facility in Chalk River, Ontario, revealed a heavy water leak that has kept the facility closed. The Ontario plant and a reactor in Petten, the Netherlands, itself set to close in July for a month's maintenance, produce the majority of the world's medical isotope supply.

On 19 June, Canada's natural resources minister Lisa Raitt appointed an expert panel to report by 30 November on how the country could ensure a stable long-term medical-isotope supply.

Three days earlier, US medical and nuclear non-proliferation groups wrote to Congress urging the United States to start its own domestic production of the isotopes, using low-enriched uranium.

P. RYAN

Safety issues have plagued the National Research Universal reactor at Chalk River Laboratories.

Climate projections taken to finest detail

Stretching the predictive capabilities of climate models, researchers last week released an unprecedentedly detailed portrait of the impacts of long-term climate change on the United Kingdom.

The projections, produced by scientists at the Met Office and the Department for Environment, Food and Rural Affairs, run to 2080. They map climate effects to a resolution of 25 square kilometres, and resolve some weather patterns down to a scale of 5 square kilometres.

"Current climate science might support projections on a scale of a couple of thousand kilometres, but anything smaller than that is uncharted territory," says University of Oxford climatologist Myles Allen, who was part of a review committee commissioned to check the report's methodology.

Regional projections were also recently produced for the United States at a resolution of 25 square kilometres, at decadal timescales up to 2039 (see *Nature* doi:10.1038/news.2009.518; 2009).

For a longer version of this story, see http://tinyurl.com/projclim

EU legislation increases clinical-trial workloads

A review of the effects of the legislation governing clinical trials in the European Union has backed up some researchers' complaints about the regulations.

The 2001 European Clinical Trials Directive, which came into force from 2004, was intended to harmonize the rules governing clinical trials. But it has faced criticism from researchers who think it is overly complex and bureaucratic (see *Nature* doi:10.1038/news.2009.163; 2009).

Last week, the final report from the European Commission-funded project assessing the directive concluded that the measures increase workloads on trial sponsors and ethics committees, and delay the start of trials. The number of commercially-sponsored clinical trials in Europe increased by about 30% between 2003 and 2007, the report adds, whereas the number of non-commercially sponsored clinical trials remained constant.

Journal of the century

With a new impact factor of 31.434, up from 28.103 last year, *Nature* maintains its position as the leading weekly science journal, according to the 2008 Journal Citation Report. Also, *Nature* was last week named 'journal of the century' by the BioMedical & Life Sciences Division (DBIO) of the Special Libraries Association (SLA).

Voted for by DBIO's 686 members, the award was given to celebrate the SLA's centenary and recognizes the most influential journal of the past 100 years.

Breaking the convention?

ILLUSTRATIONS BY D. SIMONDS

Blogs and Twitter are opening up meetings to those not actually there. Does that mean too much access to science in the raw, asks Geoff Brumfiel.

Last July, Lars Jensen carried a small shoulder bag of equipment into the atrium of the glass-and-steel conference centre in Toronto, Ontario. Jensen, a bioinformatician at the University of Copenhagen, was one of about 1,400 researchers at the annual Intelligent Systems for Molecular Biology meeting. With him he had the tools of any modern conference attendee: a laptop, a handheld PDA and a digital camera to snap a few photos of his trip.

Jensen immediately did what most researchers do: he logged on to the wireless network. He used his PDA to check FriendFeed — an online social network similar to Facebook that is popular among biologists. "Anyone else signed up for the 'orienteering' event today?" Shirley Wu, a graduate student at Stanford University, California, had written on a page that members had already devoted to the conference. "Nah," Jensen typed, "bloggers don't need icebreaker events ;-)"

Jensen's joke about face-to-face contact proved prescient. Over the next few days, he and nearly 30 other researchers met mostly via their screens in the FriendFeed conference group. During sessions, many group members posted brief comments sent from their laptops or mobile phones to the popular website Twitter, and automatically cross-posted to FriendFeed. Some of these communiqués described a comment from a talk or the flavour of a session. Other posts were links to relevant papers, or photos from around the conference centre. At one point the group even served as a public address system: "HL33: Session Chair MIA — if anyone sees Yanay Ofran they may want to point him to the session," wrote Shannon McWeeney, a bioinformatics researcher at the Oregon Health & Science University in Portland.

The virtual coverage "was not something that we had arranged beforehand", Jensen says. But by the time he and others were finished, hundreds of comments had been posted to the group. The information was so complete that Jensen, Wu and the other ring leaders were able to use it to write an authoritative conference summary that was later published (N. Saunders *et al. PLoS Comp. Biol.* **5,** e1000263; 2009).

For technophiles and advocates of scientific openness, this is the way of the future. Online groups allow meeting attendees to post and discuss research as it is presented, and follow parallel sessions. They also provide an opportunity for researchers not at the meeting, as well as a far wider community, to actively participate in it. "I think it is extremely efficient," says Jean-Claude Bradley, a chemist at Drexel University in Philadelphia, Pennsylvania. Jonathan Eisen, an evolutionary biologist at the University of California, Davis, adds that 'twittering' about a presentation he is listening to helps him to focus. "I don't want to screw up and say something that's technically wrong or conceptually wrong," he says.

> "I could take pictures of every slide and it would be on the Internet within seconds."
> — Lars Jensen

But some worry that these tools will undermine meetings. By disseminating scientific results far beyond the lecture hall, blogging and social networking blurs the line between journalists and researchers. Scientists in competitive fields may be more reluctant to discuss new findings if they can be posted on the Internet within seconds. And at a time when many conference attendees are already surfing the web rather than paying attention to the presenter, messaging is yet another annoyance. "Frankly, it can be a distraction if people are typing on their keyboards in the meeting," says David Stewart, the director of meetings and courses at Cold Spring Harbor Laboratory in New York.

Blogging without leave

Last month, Stewart unwittingly found himself at the centre of the debate about such technologies at Cold Spring Harbor's annual Biology of Genomes meeting. The meeting was oversubscribed, and many researchers were interested in following the proceedings on the web. Cold Spring Harbor streamed video of the talks online for those who had registered and paid a fee, but an informal group of bloggers also began following events.

Among them was Daniel MacArthur, a geneticist and author of the blog Genetic Future. MacArthur says that he decided to blog from the conference at the request of a number of people who couldn't attend. "They asked whether it would be possible to communicate interesting things during the sessions."

MacArthur's comprehensive postings were read by many scientists but they irked journalists

attending the meeting. The meeting rules stated that reporters had to seek permission from speakers before publishing material on their work, rules that Cold Spring Harbor instituted in part because some journals, such as *Nature*, discourage scientists from talking to the press before their work is published. But those rules didn't apply to scientist-bloggers like MacArthur and, after he posted details from a few talks, reporters contacted Stewart for clarification on the policies. The complaint was a wake-up call: "For the first time, I became aware that people were blogging about the data at the meeting," Stewart says.

Blogging can create much thornier issues for researchers. Many presenters are already cautious about revealing unpublished results at meetings for fear that rivals in the audience might note them down. Now that the note-taking is taking place live and on the web, the speed and distance that information spreads has jumped to a new level. "With the set-up I have now, I would be able to sit in a conference, take pictures of every slide that is being shown, and it would be on the Internet within seconds, while the talk is still going on," says Jensen.

This kind of direct-to-web exposure creates problems for many industrial and applied researchers. In the United States, patent applications must be filed within a year of any information becoming available to the public. The exact date of that 'public disclosure' used to be difficult to nail down, but no more, says Michael Natan, chief executive officer of Oxonica Materials, a nanotechnology company in Mountain View, California. In the Internet age, time-stamped photographs of a talk can let competitors know the exact minute a researcher presented a patentable result. Consequently, "people in industry will be much more circumspect about what they present in public", he says.

Even basic researchers have reason to fret. Last year, a group of theoretical physicists photographed slides from a meeting presentation, extracted the data, and used them in their own analysis, which they published online (see *Nature* **455,** 7; 2008). In that case, the theorists were given permission by a presenter, and the photographs were properly cited, but the situation illustrates how easily data can find their way into the public sphere. In many fields, competition is so intense that you must conceal to survive, says Natan.

"Frankly, it can be a distraction if people are typing on their keyboards in the meeting."
— David Stewart

For denizens of the blogosphere, these sorts of concerns seem a little out of date. "I think scientific conferences are about your sharing with the world what you're doing," says Francis Ouellette, a researcher at the Ontario Institute for Cancer Research in Toronto, who twittered at the Cold Spring Harbor Meeting. "Whether or not the participant you're sharing with is in the room is somewhat inconsequential."

Ouellette and many other active bloggers are also members of the 'open science' movement, which encourages researchers to make their data public as quickly as possible. Bradley sees this openness as a powerful deterrent to anyone hoping to scoop him at a conference because anything cribbed from his talk is already out on the Internet for everyone else to view. "If someone actually does copy something, I think it would be pretty embarrassing," he says, "it's already there, and it's indexed to Google."

The rights of scientists

Prompted by his recent experience, Stewart has come up with a pragmatic solution. At future meetings, anyone communicating information to third parties, whether by news story, blog or 'tweet', will now be required to ask presenters beforehand. "What we really want to do is to protect the rights of individual scientists to present their data in a pre-published form," he says. "I'm not saying don't blog, I'm saying blog by all means, but get permission." MacArthur says that the new policies seem like a fairly good compromise, but, he adds: "I'm hopeful that other conferences will tend to adopt more open policies."

Conference organizers contacted by *Nature* had a wide range of policies on social networking. Many societies have banned digital photography in talks and poster sessions and some consider bloggers to be members of the media and subject them to certain reporting restrictions. However, almost nobody has developed a policy on when twittering is fair play. "This has not come up in the past but it may be something we consider in the future," says Kevin Wilson, a spokesman for the American Society for Cell Biology in Bethesda, Maryland.

Journals are also pondering how best to handle social networking at meetings. *Nature* generally supports social media tools, says Philip Campbell, *Nature*'s editor-in-chief. And as long as it's not a deliberate attempt to hype a new finding, he says that researchers should feel free to talk to colleagues who blog or twitter. Ginger Pinholster, the director of public programmes for the American Association for the Advancement of Science, the publisher of *Science*, agrees. As long as the scientist wasn't trying to promote his or her work to the public, it wouldn't be a problem. She adds that blogging unpublished results is a problem that "we just haven't run into yet".

Although Cold Spring Harbor has opted for control, the organizers of the Intelligent Systems for Molecular Biology meeting are choosing total openness. When this year's meeting opens in Stockholm later this week, they are planning to fully embrace social networking tools. FriendFeed entries will be created for each talk at the meeting, and the entries for the keynote sessions will be posted directly to the meeting's main website. That means some people could get the buzz of the meeting without travelling to Stockholm at all.

But that's not why Lars Jensen is staying at home this year. The main papers being presented have already been published, he says, so there won't be much new. He might follow the sessions online, but then again, he might catch up the old-fashioned way. "I have colleagues who are going," he says. "I can always ask them afterwards whether anything interesting happened." ■

Geoff Brumfiel is a senior reporter for *Nature* in London.

See Editorial, page 1033, Essays, pages 1054–1057, and online at http://tinyurl.com/sciencejournalism.

CORRESPONDENCE

Interrogation: hard for psychologists to act as whistleblowers

SIR — I was disappointed in your Editorial 'Responsible interrogation' (*Nature* **459,** 300; 2009). You accept the role of psychologists in secret interrogations because they can supposedly act as whistleblowers against inhumane treatment. But your argument ignores nearly fifty years of psychology research, beginning with the classic work of Stanley Milgram (see, for example, *J. Abnorm. Soc. Psychol.* **67,** 371–378; 1963), showing that isolated individuals — including, yes, even psychologists — are notoriously poor at resisting pressure from authorities to act in abusive ways.

The position you take is naive, and in adopting it you allow yourself to be 'played' by those elements in the government and the military who are most willing to commit extreme violations of human rights.

Michael R. Jackson Department of Psychology, Earlham College, 801 National Road West, Richmond, Indiana 47374, USA
e-mail: jacksmi@earlham.edu

Interrogation: our professional body forbids involvement

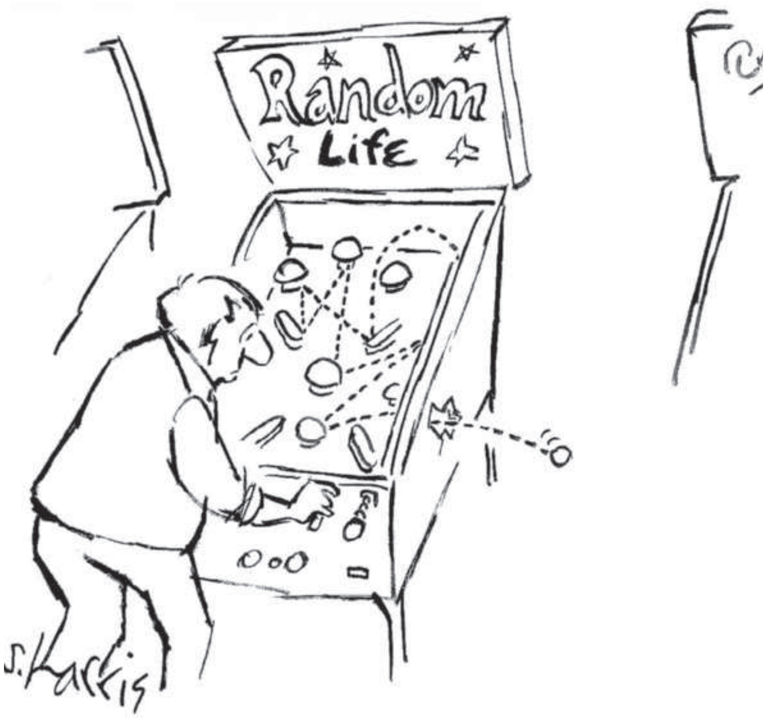
SIR — Your Editorial 'Responsible interrogation' (*Nature* **469,** 300; 2009) takes a remarkably unscientific approach to the topic of psychologists' participation in national-security interrogations, in that it omits important current facts. The scale of this omission is comparable to failing to mention *Helicobacter pylori* in a discussion of peptic ulcers, the treatment of which was revolutionized by the discovery that most are caused by *H. pylori*.

Psychology's attitude to the ethics of participating in national-security interrogations was similarly revolutionized by a petition resolution passed by the American Psychological Association (APA) membership last year (see http://tinyurl.com/lb9nd5). This has been adopted by APA's governing body under the name of 'Psychologists and unlawful detention settings with a focus on national security'.

As a result of this resolution, and contrary to the implications of your Editorial, psychologists may not participate in national-security interrogations. The resolution, which constitutes current APA policy, states: "Be it resolved that psychologists may not work in settings where persons are held outside of, or in violation of, either International Law (*e.g.* the UN Convention Against Torture and the Geneva Conventions) or the US Constitution (where appropriate), unless they are working directly for the persons being detained or for an independent third party working to protect human rights."

Laurel Bass Wagner Division of Psychoanalysis, American Psychological Association, 6060 N. Central Expressway, Suite 332, Dallas, Texas 75206, USA
e-mail: lbwagner@tx.rr.com

Readers are welcome to comment at http://tinyurl.com/mc3byy



Free will: it's a normal biological property, not a gift or a mystery

SIR — In his Essay 'Is free will an illusion?' (*Nature* **459,** 164–165; 2009), Martin Heisenberg argues that humans must have free will because freedom of action has been demonstrated in other animals — including those as small as fruitflies and bacteria.

Heisenberg's case rests on a combination of random chance and lawfulness, escaping the classic two-horned dilemma of determinism versus indeterminism that is so popular in introductory philosophy courses and textbooks. Starting with William James in 1884, such a two-stage combination of 'free' and then 'will' has frequently been proposed by philosophers and scientists, notably by some quantum physicists after Martin's father, Werner Heisenberg, established irreducible physical randomness with his indeterminacy principle in 1927. But academic philosophers, particularly those who work in the Anglo-American school of analytical language philosophy, have been reluctant to embrace these ideas.

The philosophers' standard argument against free will is simple and logical. If our actions are determined, we are not free. If nature is not determined, then indeterminism is true. Indeterminism implies that our actions are random. If our actions are random, we did not will them.

Heisenberg's proposal makes freedom a normal biological property of most living things, and not a metaphysical mystery or a gift from God to humanity. The genius of this proposal is that it combines randomness with an adequate macroscopic determinism consistent with microscopic quantum mechanics.

John Locke wrote in his *Essay Concerning Human Understanding, Book II,* that it is not the will that is free but the man. The will determines our actions. Heisenberg writes that Kant would have been pleased. Locke too might have been pleased to see this return to common sense. We may not have metaphysical free will but we do have biophysical free will.

Robert O. Doyle Astronomy Department, Harvard University, 77 Huron Avenue, Cambridge, Massachusetts 02138, USA
e-mail: bobdoyle@informationphilosopher.com

Free will: emotions and consciousness could contribute

SIR — In his Essay 'Is free will an illusion?' (*Nature* **459,** 164–165; 2009), Martin Heisenberg suggests that belief in free will is supported by quantum events.

"Science journalists don't get to witness earlier drafts of history-making because these are part of the peer-review process." Toby Murcott, page 1054

However, the concept of free will may become confused if it is linked with an absence of determinism.

As an example, let us consider three schoolgirls, X, Y and Z, confronted with the proof of Pythagoras's theorem: X has a talent for mathematics and enjoys working out proofs; Y is weak in this domain and is unquestioning; Z has average ability but her decisions are capricious. The teacher instructs them to believe the theorem because it is correct. Y accepts it immediately, X first confirms for herself that the proof is valid, but Z (possibly influenced by a 'quantum event' in her brain) refuses to agree. Although the behaviour of X and Y is predictable and determined, given their personalities and abilities, Z's is not.

Heisenberg's suggestion would support the conclusion that only Z's decision was 'free'. But X could be judged as the one who made the really free (autonomous) decision. Y's decision is formally free, having been determined by her accepting nature, but it is undermined because it stems from the teacher's authority. Z's reaction is not free at all, because it was not determined by Z herself but by a random event in one or more of her brain cells.

In short, deciding freely does not imply a lack of determinism — rather, it is determined by central aspects of our personality: our long-term needs, the emotions accompanying their non-fulfilment, and our rational thinking about the means to satisfy those needs. Our decisions may therefore not be completely free, because they are not always exclusively determined by these central (core) factors. A person who stops smoking on rational grounds is freer than another who makes a decision to stop but fails to do so.

Quantum events have no relevance here: the question is whether we are influenced more by our core factors than by drives that are not rationally founded, such as habit, addiction or external pressure. Consciousness and the experience of positive or negative emotions could well play a part in our decisions: in my opinion, these are not epiphenomena — mere parallel events — but essential for bringing about determining factors that underlie our free will. This would not exclude a purely naturalistic explanation of the processes that we experience as consciousness.

Etienne Vermeersch University of Ghent, Blandijnberg 2, 9000 Ghent, Belgium
e-mail: etienne.vermeersch@ugent.be

Readers are welcome to comment at http://tinyurl.com/m2ybmo

How air capture could help to promote a Copenhagen solution

SIR — Your News Feature 'Sucking it up' (*Nature* **458,** 1094–1097; 2009) reports on the issue of the capture of carbon dioxide from air. This is timely, as in February this year, President Obama and the Canadian Prime Minister, Stephen Harper, agreed to work together on carbon capture and sequestration as part of an effort to build a North American environmental and energy accord. US and Canadian government funding for 'carbon capture and storage' (CCS) projects has ballooned during the past six months, and in May the US energy secretary, Steven Chu, announced CCS funding of $2.4 billion, which specified for the first time "CO_2 capture from the atmosphere".

Conventional CCS has been used successfully since 1996, but it has many critics. It has been blocked in global climate negotiations and is likely to be a contentious issue at meetings of COP15 — the conference of the United Nations Framework Convention on Climate Change, to be held in Copenhagen in December this year — which will decide on the future of the Kyoto Protocol after 2012.

Opposition to CCS, and support for the attendant 'cleaner coal' approach, is motivated by incentives to continue using fossil fuels, rather than making the transition to renewable sources of energy, and by the effort needed to retrofit and clean up existing fossil-fuel plants, which are responsible for more than 40% of global emissions. In new power plants, conventional CCS can at best neutralize carbon emissions.

Air capture could satisfy these critics, as well as potentially strengthening the president's proposal. The technology is under evaluation by the American Physical Society and is rapidly gaining support in the business community. It will encourage nations to cooperate at global negotiations, including China and developing countries, because the ubiquity of air means that this technology can be used by everyone; small emitters such as Latin America and Africa will be able to decrease their atmospheric carbon beyond what they actually emit. When driven by renewable energy, air capture will help the transition to renewable energy. Incorporating air capture into the Clean Development Mechanism of the Kyoto Protocol would be a big step forwards.

Two footnotes to your News Feature are in order. First, G.C. is co-inventor and co-owner of the air-capture company Global Thermostat's technology. Second, P.E.'s mention of how much it might cost on a global scale to reduce CO_2 is only an estimate — specific costs for the company's technology are awaiting data from the commercial demonstration plant that is to be built in the near future.

Graciela Chichilnisky Departments of Economics and Statistics, and Columbia Consortium for Risk Management, Columbia University, New York, New York 10027, USA
e-mail: chichilnisky1@gmail.com
Peter Eisenberger Department of Earth and Environmental Sciences, Columbia University, New York, New York 10027, USA

Planck's power lies in its unique instrument combination

SIR — Your News Feature 'The test of inflation' (*Nature* **458,** 820–824; 2009) highlights some of the exciting scientific data to be collected by the European Space Agency's Planck spacecraft, and the theoretical issues underlying its objectives to study the moments after the Big Bang.

You discuss the new high-frequency 52-bolometer detector, but do not mention the 22-radiometer low-frequency instrument (LFI). However, it is this unique combination of Planck radiometers and bolometers in an integrated focal assembly that is key to achieving the broad spectral coverage needed to separate foreground emission of galactic and extragalactic origin from the cosmological signal.

This feature is essential in searching for subtle signatures in the cosmic microwave background, including possible clues from an inflationary era of the Universe (such as polarization B-modes and non-Gaussianity). Planck's 70-gigahertz radiometers will observe the sky in the frequency band that is least contaminated by foregrounds, and with a sensitivity and angular resolution surpassing that of all previous experiments.

Following the 1989 Cosmic Background Explorer differential microwave radiometer and the 2001 Wilkinson microwave anisotropy probe, the Planck radiometric instrument will use polarization-sensitive detectors based on indium phosphide HEMT (high electron-mobility transistor) cryogenic amplifiers cooled to 20 K. It is the data from these detectors, combined with the bolometer data, that will give Planck its superiority over its predecessors.

Nazzareno Mandolesi and Planck LFI co-investigators IASFBO – INAF, Via Gobetti 101, 40126 Bologna, Italy
e-mail: mandolesi@iasfbo.inaf.it

ESSAY

Science journalism: Toppling the priesthood

In the first of three essays, Toby Murcott argues that the process of science needs to be opened up if journalists are to provide proper critique.

There is a rhythm to science news, easy to spot in the mainstream media and as familiar to every science journalist as breathing. It follows the publication cycles of the major peer-reviewed journals such as *Science*, *The Lancet* and *Nature*. As press releases describing research arrive in our inboxes they are scanned for stories, the most newsworthy picked, offered to editors and then reported.

This is not unusual. As British journalist Nick Davies points out in his book *Flat Earth News*, much of the news agenda in all fields is press-release driven. Journalists are of course trained to stand back and provide a critique, including context and a broader perspective, rather than simply reporting what they read in a press release. But doing so is a particular challenge for science journalists.

To best serve our audiences, we journalists need to be able to see how a new finding fits into the field, know when something new is significant, and have the knowledge and the confidence to ask searching questions. I have a PhD in biochemistry and three years postdoctoral research, so if I am reporting a discovery in my field, I can make a reasonable attempt at understanding the technical detail and will have a sense of the overall history, evolution of ideas and current debates. I will know who is a leader in the field, and who is an outlier; I will be able to distinguish majority views from minority ones. Yet as a science journalist I am expected to cover more than just biochemistry. I need to be able to report on findings in cosmology, ecology, particle physics and much more. To draw on the knowledge of scientists in these fields, I must first find out which scientists are most relevant, and have a sense of their opinions and place within the field. All of this takes time, which reporters often don't have.

This is partly why so many journalists resort to doing the bare minimum: reproducing press releases. Many journalists will telephone or e-mail one of the main contributors given on the press release to ask a few supplementary questions; but there is rarely the time or the expertise to go into the full story of how an item of research came to be, and how it fits into the bigger picture.

Other (non-science) journalists recognize this and are uncomfortable with what they see. As science correspondent for the BBC World Service, I regularly experienced the quiet frustration some elements of the newsroom felt with science journalists. My colleagues felt that we reported on published papers without significant analysis, depth or critical comment: we just translated what scientists said.

ILLUSTRATIONS BY J. FIELD

The priest perception

You could say that this is not exactly a description of a journalist — more that of a priest, taking information from a source of authority and communicating it to the congregation.

This perception is reinforced when you compare our role with that of other journalists. Political journalists, for example, take an active part in the political debate. They produce expert commentary on the subtleties of the political process, highlighting strengths, weaknesses and potential pitfalls of policy ideas. They interview politicians as equals, challenging them to explain their ideas and, crucially, picking them up on inconsistencies, contradictions and mistakes.

These journalists are active participants in the process of knowledge creation in a way that science journalists cannot be, given the qualifications needed to act as an equal in scientific debate. Although science news reporting can influence science funding and research priorities, science journalists are not players in the scientific process. Again this is like a priest, who has little or no effect on the activities of the deity itself and who is not actually needed for the deity to continue.

> "Greater public awareness of the peer-review process would benefit us all."

The 'priesthood' model of science journalism needs to be toppled, but this is easier said than done. The time pressures on journalists today do not bode well for calls for more depth, context and criticism. But one appealing way to start is if scientists helped to unmask the very human process through which science is produced and reviewed, thus dismantling their church-like roles as unquestionable authorities. Press officers at universities and research journals could help by providing more background and context to new findings and discoveries for those journalists who are still too pressed for time to obtain this information themselves.

Alongside this is a need for science journalism to develop roles analogous to those of political journalism or literary and artistic criticism. We need to have the willingness to acquire more expertise so we can understand the technical details of the science, be able to interrogate and be critical when necessary, and not feel intimidated by those we are interviewing.

Crafting knowledge

All over the world, scientists, educators and policy-makers rightly call for more public understanding — or public awareness — of science. But many assume that this means more media coverage of the latest genome sequence, or the latest Hubble image. And journalists oblige by serving up stories along these lines. Genuine public awareness of science, however, also includes an understanding of how scientific knowledge is crafted. This is something that scientists and editors of research journals know intimately, and often talk about. The broader public deserves to know too.

Journalism is often described as history's

first draft. Much contemporary science journalism, however, can be seen as a second, or even a third draft. Unlike reporters covering other fields of public life, science journalists don't get to witness earlier drafts of history-making because these are part of the peer-review process.

One way to tackle this would be to allow journalists access to (anonymous) referee comments alongside a final paper. This would have implications for how journals communicate their findings, and would need more discussion among the science community before being broadly implemented (some journals, such as *Biology Direct*, already publish reviewer's comments alongside their final papers). On balance, I think it would be beneficial to science communication if this were the usual practice.

When we read a published scientific paper, it often feels like the last word, even though there is often a gripping story behind and ahead of it. Unless they are very experienced, journalists will have little sense of these stories, and that what they are reading has in fact been through an extensive review process of critique, comment, discussion, argument and refinement. Access to that process would help to make stories richer and more compelling for audiences. It is also likely to reduce the tendency — or should I say temptation — to report every new finding as a 'breakthrough', something that happens all too often in the general media.

Greater public awareness of the peer-review process is something that we will all benefit from. It will help journalists do a more thorough job of reporting science; and it will aid the public understanding of science. ■

Toby Murcott is a writer and broadcaster and former science correspondent for the BBC World Service. A longer version of this Essay will appear in *Communicating Biological Sciences*, which will be published in November by Ashgate.
e-mail: toby.murcott@ketoe.co.uk.

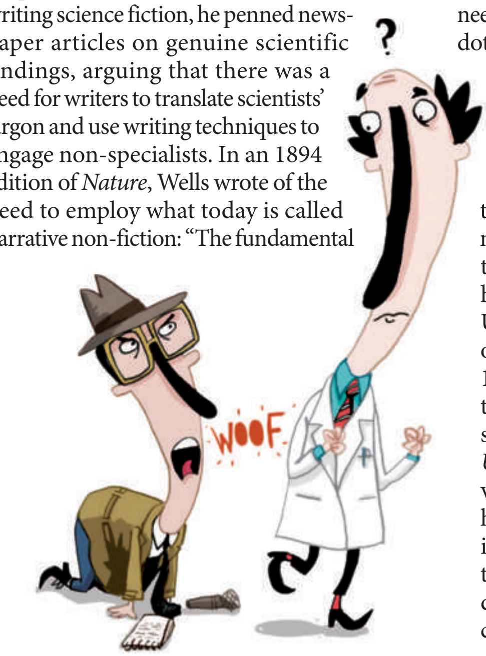
See Editorial, page 1033, and online at http://tinyurl.com/sciencejournalism.

Science journalism: Too close for comfort

In the second of three essays, Boyce Rensberger tracks the progression of scientific correspondents from cheerleaders to watchdogs.

Science journalism has undergone profound changes since its origin more than a century ago, some more radical than most journalists of today are aware. Although there are legitimate complaints that some current reporters are too close to their sources, or otherwise unable to deliver a disinterested analysis of the field, it is salutary to reflect on how far the profession has come since its beginning.

In the 1890s, there seem to have been no full-time science journalists in either the United States or Britain, although there was one notable part-timer — H. G. Wells. When he wasn't writing science fiction, he penned newspaper articles on genuine scientific findings, arguing that there was a need for writers to translate scientists' jargon and use writing techniques to engage non-specialists. In an 1894 edition of *Nature*, Wells wrote of the need to employ what today is called narrative non-fiction: "The fundamental principles of construction that underlie such stories as Poe's 'Murders in the Rue Morgue', or Conan Doyle's 'Sherlock Holmes' series, are precisely those that should guide a scientific writer." (See *Nature* **50,** 300–301; 1894.)



In 1904 Adolph Ochs, founder of the modern *New York Times*, hired the legendary Carr Van Anda as his managing editor. Van Anda may have been the most scientifically astute news executive of the twentieth century. He had studied astronomy and physics at university, wrote science stories and encouraged his reporters to cover science. He stressed the need for accuracy: in an often-quoted anecdote, Van Anda corrected a mathematical error in a lecture of Albert Einstein's that *The New York Times* was about to print — after, of course, checking with Einstein.

Unlike now, the work of most science reporters in those days consisted largely of translating jargon and explaining the statements of scientists and medical leaders. More than that, according to Bruce Lewenstein, a historian of science journalism at Cornell University in Ithaca, New York, the handful of science journalists at US newspapers in the 1930s and 1940s believed that it was their job to persuade the public to accept science as the salvation of society (B. V. Lewenstein *Public Underst. Sci.* **1,** 45–68; 1992). This was a vestige of the Progressive Era in American history that spanned the 1890s to the 1920s, in which intellectuals of all stripes believed that society was perfectible and that the wonders of science and technology would lead civilization towards this ideal.

To do this job, US science reporters decided that they needed scientists to take them more seriously and thus created the National Association of Science Writers (NASW) in 1934. They explicitly called themselves 'writers' rather than 'journalists' or 'reporters' because they felt it sounded more professional. NASW members, all 12 of them, told scientists that they could feel safe talking to members because they belonged to an elite society. Scientists were advised not to talk to non-members, because those reporters were not 'true science writers'. Thus began what I call the 'Gee-Whiz Age' of science reporting, in which the emphasis was on the wonders of science and respect for scientists, rather than on any analysis of the work being done or any anticipation of its effects on society.

In 1937, George W. Gray, who covered science for *Time* magazine, wrote that science journalism should serve to make "the scientific method an integral part of popular education and through it a universal element of

ILLUSTRATIONS BY J. FIELD

NATURE|Vol 459|25 June 2009

civilized thinking". When Gray won an award in 1949 from the American Association for the Advancement of Science, he gushed: "What counts most is the recognition from scientists themselves."

Throughout much of this period science reporters encouraged the mythos, as Lewenstein puts it, "that the proper relationship between scientists and science writers was one of trust and respect." One of the best examples of such trust and respect can be seen in science reporter William Laurence's coverage of the Manhattan project to develop the first atomic bomb.

Laurence wrote for *The New York Times* and was well known for his desire that the public appreciate science — so much so that the Pentagon invited him to be the only reporter privy to the project and the atomic bombing of Japan. He was, in effect, one of the first science journalists to be embedded within the military. As it happens, Laurence was also on the US government's payroll, writing press releases about the bomb for President Harry S. Truman and the war department. Incredibly, Laurence's government job was not kept secret, nor did it seem to strike anyone as the massive conflict of interest that journalists today would see.


F. GORO/TIME LIFE PICTURES/GETTY

William Laurence (left) wrote press releases for the US war department even while reporting on J. Robert Oppenheimer's (right) atomic bomb.

Friendly fire

After Hiroshima and Nagasaki, as rumours developed that people outside the blast area were becoming ill and dying, the war department held a press conference at the military base at Alamogordo in New Mexico. Laurence opened his resulting story this way: "This historic ground in New Mexico, scene of the first atomic explosion on earth and cradle of a new era in civilization, gave the most effective answer today to Japanese propaganda that radiations [sic] were responsible for deaths even after the day of the explosion." He wrote that the tour of Alamogordo would "give the lie to these claims". The journalism community felt that Laurence epitomized the best in science journalism: in 1946 he was awarded the Pulitzer Prize.

In the 1950s, science journalists wrote often of their belief that the facts of scientific discovery — the joy of science for its own sake — should be emphasized over the practical and social implications. NASW members would talk of how to make science journalism 'more effective', by which they meant making it result in the wider public appreciation of science, rather than making it a better watchdog of the field. It was this attitude, I believe, that led many news executives in those days to begin to distrust their own science and medical writers. After all, in every other part of the newsroom, reporters are valued for their disinterested, even aggressive stance towards the people they cover.

Cracks soon began to appear in science journalism's 'Gee-Whiz' zeitgeist. In 1962, Rachel Carson published her book *Silent Spring*, highlighting the role of the pesticide DDT in weakening birds' eggshells and killing wildlife. But because pesticides had previously been reported as technological marvels, something the public should admire and accept, science journalists were conflicted. Some lambasted Carson's book as emotional and biased. Lawrence Lessing, who wrote for *Time* and *Fortune* magazines, attacked her and said that the best science writing was a result of cooperation with the American Chemical Society and with "enlightened industries encouraging a better discourse between working scientists and science writers, to the benefit of the public". Yet Carson's work — and her view that industry claims should not be taken uncritically — helped to launch not just the environmentalist movement, but also the field of environmental reporting. These reporters were more willing to be critical of the work of scientists, especially those funded by industry.

The 1970s offered increasing evidence of technology's potentially adverse effects, in part owing to controversies and crises such as the reactor meltdown at the Three Mile Island nuclear power plant near Harrisburg, Pennsylvania. By this time there was no way science journalists could ignore the social and political implications of their topic. And so the next great age of science journalism began — the 'Watchdog Age' — as science reporters became much more like their colleagues in other parts of the newsroom.

Paper trail

The quantity of science journalism boomed too, starting with the birth of a science section in *The New York Times* in 1978. By the boom's peak in 1987, according to one count, some 147 newspapers had at least a weekly science page, and four new popular-science magazines had joined the venerable *Scientific American* and *Science News*. Sadly, this upturn was short-lived. By the late 1980s the magazine upsurge was over — all the new titles, none of which was ever profitable, died except for *Discover*, which still exists alongside the older titles. The number of newspaper science sections started to fall, eliminating all but the handful that survives today.

We are obviously now in the 'Digital Age', and the very definition of journalism is changing in uncertain directions. Science journalism has moved from working for the glory of the scientific establishment to taking back its independence and exercising a new responsibility to the public. Now, traditional news outlets are withering, leaving many journalists to self-publish online with total independence and a direct connection to the public. But scientists too can use the web, bypassing journalists altogether and taking their science — and their agendas — directly to the public. It is becoming increasingly difficult for readers to tell which sources are disinterested and which have an axe to grind.

If science journalists are to regain relevance to society, not only must they master the new media, they must learn enough science to analyse and interpret the findings — including the motives of the funders. And, as if that were not enough, they must also anticipate the social impacts of potential new technologies while there is still time to make a difference. ■

Boyce Rensberger was director of the Knight Science Journalism Fellowship programme at the Massachusetts Institute of Technology in Cambridge from 1998 to 2008, and a science reporter for 32 years, chiefly at *The Washington Post* and *The New York Times*.
e-mail: boycerensberger@gmail.com
See Editorial, page 1033, and online at http://tinyurl.com/sciencejournalism.

Science journalism: The Arab boom

As research increases in Arab countries, the media is stepping up to report on it. In the third of three essays, **Nadia El-Awady** says the local journalism has much room for improvement.

N. EL-AWADY

Science gets a good quantity of coverage in Arab papers, but is it good quality too?

When I started covering science for IslamOnline.net in Egypt about ten years ago, I did an Internet search for 'science journalism' to see if such a profession even existed. I had not heard the two words put together before, at least not in an Arab context.

Things have changed. In 2006 I helped to launch the Arab Science Journalists Association (ASJA), which has since grown to 179 members as of May 2009. We are now co-bidding — with the US National Association of Science Writers — to host the 7th World Conference of Science Journalists in Cairo in 2011 (the 6th conference is being held in London next week). But as Arab science journalism gains strength, we need to ensure that its quality grows in step with its quantity.

Although the science staff of media organizations in the United States and Europe face cutbacks, a survey of ASJA members in January 2009 indicated that full-time jobs for Arab science journalists have remained relatively stable over the past five years, and freelancers report more opportunities. The *Al-Ahram* daily newspaper in Egypt, for example, which started a regular weekly science section in the 1950s, employs 20 full-time science journalists. *The Washington Post*, a US paper with a similar circulation, has eight. Al Jazeera, despite being one of the largest and most famous Arab news outlets, has no full-time science reporters, but Egypt's Al-Manara science research channel has 40 full-time staff. This promising situation is partly attributable to a heightened interest in science by Arab audiences, and a growing amount of scientific research and conferences in the region.

As local research is growing, Arab science-related institutions are showing greater interest in supporting media activities, in part to promote science and its communication. For instance, the Arab Science and Technology Foundation, based in Sharjah in the United Arab Emirates, finances many ASJA activities. And Egypt's Academy of Scientific Research and Technology funds the local *Al-Ilm* science magazine and Al-Manara's science channel. The funding and the acknowledgement of the importance of media are welcome, but there is need for caution. Journalists must be able to critically cover the institutions that fund them.

There is a deep reverence for scientists and doctors in Islamic cultures, so we need be careful that science journalists do not feel they owe a debt of gratitude to the very people that they should be investigating. Many science writers are also working scientists — one-fifth of ASJA's members have a faculty position, are working in research or are working doctors or veterinarians. This can make it hard to maintain a critical perspective in the journalistic work.

Unintentional secrecy

Despite these strong links between science and the media, Arab science institutions remain relatively closed to journalists. As I sit at my desk in Cairo, it is easier for me to know what is happening in American universities halfway across the globe than to know what is happening within the walls of Egypt's National Research Center just across the street. There is a lack of communications officers in Arab science-related institutions and universities. Press releases are rarely sent to journalists, and websites are often poor and badly maintained.

"Uncritical reporting of H1N1 policies by Egypt's media led to unnecessary panic."

Another problem for science journalism stems from a more general issue. Many media platforms are government-owned and, as a result, many journalists provide uncritical coverage of governmental announcements. The results can be seen, for example, in Egypt's recent coverage of the H1N1 'swine flu' outbreak. The country's pig population was slaughtered by presidential decree — despite no evidence of the virus in the pigs. Before a single H1N1 case had hit the country, to help limit infection Egypt's Minister of Health, Hatem El-Gabaly, called for prayers and university exams to be held in open spaces rather than in mosques or enclosed rooms. Egyptian newspapers reported that the Ministry of Agriculture had banned all imports of pig gut and hair used in the manufacture of surgical threads and shaving brushes. Uncritical reporting of these policies by Egypt's media led to unnecessary panic among the general population.

Other factors compound these issues. A lack of English-language skills, combined with a lack of scientific sources in Arabic, limits available information and participation in international events. Some journalists and scientists in the region tend to consider the Internet in and of itself as a source of credible information.

Arab science journalists need to address these issues. They must diversify their funding and be mindful of the importance of distancing themselves from funding sources. English-language skills need enhancement and more training in Arabic should be available. Journalists and scientists need training on what are credible sources of information.

If we don't nurture and sustain the rise of Arab science journalism, and if we leave our weaknesses as they are, then that rise may be the harbinger of its downfall. ■

Nadia El-Awady is a freelance science journalist based in Cairo. She is a founding member and past president of the Arab Science Journalists Association and a board member of the World Federation of Science Journalists.
e-mail: nadiaelawady@gmail.com

See Books & Arts, page 1061, and online at http://tinyurl.com/sciencejournalism.

BOOKS & ARTS

From rubber ducks to ocean gyres

The tale of Curtis Ebbesmeyer's use of beachcombing to reveal patterns of ocean circulation conveys the romance of early marine science, but his lessons for today are serious, finds Simon Boxall.

Flotsametrics and the Floating World: How One Man's Obsession with Runaway Sneakers and Rubber Ducks Revolutionized Ocean Science
by Curtis Ebbesmeyer and Eric Scigliano
HarperCollins: 2009. 304 pp. $26.99, £16.99

As I gaze down from the aeroplane window at the Arabian Sea, the irony that I'm flying around the world to talk about the impact of climate change and how to reduce carbon emissions is not lost on me. To get to my meeting without emitting any carbon, I could, like the ancient Norsemen, have ridden the ocean gyres to far-off lands. Although the excuse "Sorry I'm 3.4 years late but that's the Majid Gyre for you" would be original, it might not satisfy my audience. Yet it comes out of a new book on ocean currents by oceanographer Curtis Ebbesmeyer and journalist Eric Scigliano.

I say the Majid Gyre as this is what the authors unilaterally call the Indian Subtropical Gyre. And I say a precise 3.4 years because that is the period of a harmonic in the ocean's large-scale flow. *Flotsametrics and the Floating World* highlights that the 11 major gyres that determine the surface circulation of the oceans rotate with a fundamental tone of 13.4 years, with half, quarter (hence 3.4), eighth and sixteenth tones. The authors elegantly call this the 'music of the gyres'.

As I read on, however, my attention turns to my immediate surroundings. Plastic is everywhere: in the cabin lining, the chairs, the food trays and, let's face it, most of the airline food too. Virtually everything we use today is plastic, but our disposal of it is haphazard. As plastic objects break down into smaller and smaller particles, they disappear from view and from our minds but not from our environment. Most of this plastic will one day end up in the ocean.

Ebbesmeyer and Scigliano show — from journal papers, reports and their own observations — that the level of plastics building up in the sea, and particularly in the ocean gyres, is alarming. The North Pacific Subtropical Gyre has recently been dubbed the "Garbage Patch" by the media and by scientists; in places, plastic dust particles exceed phytoplankton in number.

Plastic bags are one source that has recently received a lot of attention. Some states in India have banned them, and even carrying one in the street incurs a penalty. This may be warranted. Although initial investigations of the environmental impact of ubiquitous microscopic plastic particles showed them to be benign roughage, recent research, covered in the book, flags up health issues. Chemicals from plastics can be carcinogens and 'endocrine disrupters' that mimic hormones such as oestrogen.

A. SEALE/IMAGEQUESTMARINE.COM; D. INGRAHAM

Curtis Ebbesmeyer has raised concerns that our seas are being choked by rubbish and plastic debris spilt from ships' containers.

It is not only plastics that find their way into the ocean through landfill and dumping at sea. Flotsam and jetsam also get caught in its endless gyres. Ships are not always obliged to report the loss of containers at sea, nor their contents, and many do not because it looks bad on their records. In international waters the law against such pollution is inefficient or unenforceable.

Yet Ebbesmeyer has built much of his ocean-circulation research on the serendipitous finds of beachcombers and mariners — from container loads of plastic bath toys and trainers spilled from ships, to bodies and ghost vessels lost at sea.

Flotsametrics and the Floating World is a tale of climate change and its effect on ocean gyres, and of seas so choked with plastic that climate change may be the least of our problems. The book also touches on darker, intriguing issues, such as the fate of bodies in the marine environment. But there is good news here too, from the discovery of new drugs in marine organisms to how

oceans have enabled the dispersal of human civilizations across Earth. Best of all, the book reads like a good story because of the autobiographical way in which Ebbesmeyer and Scigliano lead the reader through.

My own path through oceanography has been similar to Ebbesmeyer's. We both worked on similar problems; from slabs of water and their mixing to tracking spills, dead bodies and plastic ducks. In the 1970s and 1980s, oceanography involved more exploration and pure experimentation than is typical today — it was romantic and fun. We get more out of marine science now, using technological tools such as satellites, supercomputers, robotic drifters and acoustic probes. But we have also inherited a bureaucracy of targets, cost allocations and concerns over health and safety, and this change in culture comes through in Ebbesmeyer's story. Concern for safety is no bad thing. But picture the scene if we were to start oceanography from scratch today — "Yes, we're going out in very deep water with 10-metre waves, in a small metal container while operating high-voltage equipment and explosives." We would be banned.

Whether you want to learn more about how the oceans tick or how we are affecting our environment, or to reminisce about science not being what it used to be, this is a very enjoyable, if at times dark, book. And I never did have to resort to the in-flight movie. ■

Simon Boxall is a lecturer in oceanography at the National Oceanography Centre, University of Southampton, Southampton SO14 3ZH, UK.
e-mail: simon.boxall@noc.soton.ac.uk

Cooking debate goes off the boil

Catching Fire: How Cooking Made Us Human
by Richard Wrangham
Basic Books: 2009. 320 pp. $26.95

Richard Wrangham's new book is an expansion of a provocative hypothesis about human history that he and his colleagues first published in 1999. Cleverly evoking anthropologist Claude Lévi-Strauss's work *The Raw and the Cooked*, the article was entitled 'The Raw and the Stolen: Cooking and the Ecology of Human Origins.' Wrangham restates his hypothesis concisely and elegantly at the outset of *Catching Fire*:

"I believe the transformative moment that gave rise to the genus *Homo*, one of the great transitions in the history of life, stemmed from the control of fire and the advent of cooked meals. Cooking increased the value of our food. It changed ... our social lives. It made us into consumers of external energy and thereby created an organism with a new relationship to nature, dependent on fuel."

The primary criticism of the cooking hypothesis, as it was first published, centred on the lack of convincing evidence that cooking began about 2 million years ago in Africa. Wrangham, a biological anthropologist at Harvard University, attributes the origin of the large-brained *Homo erectus* and its geographic expansion out of Africa to the invention of a new mode of preparing tubers that greatly enhanced digestibility and nutritional benefits. The hypothesis was sceptically received in 1999; commentators used the phrases "half-baked" and "Just-so story". After an additional ten years of research, I hoped Wrangham would present a stronger argument. He does not.

The oldest credible evidence of controlled fire is around 790,000 years old, from a site at Gesher Benot Ya'aqov in Israel. Here, burned flint fragments were clustered, suggesting the use of hearths. But burned seeds and wood, also present, were not clustered near the hearths or one another. So apart from the burned flint, the evidence for controlled fire at this site is equivocal, as is all such evidence before about 250,000 years ago.

The lack of evidence is daunting. If one individual and his descendents made a fire once a week for a million years, they should have left the remains of some 50,000,000 fires. "Evidence of humans controlling fire is hard to recover from early times," Wrangham says.

The main ingredient of his hypothesis is, succinctly, potatoes. Early hominins used cooked tubers and underground storage organs of plants as an important fall-back food. Tubers were underused by other species because they require digging and cooking for maximum benefit.

But Wrangham runs afoul of the evidence here too. Stone tool assemblages from before 2 million years ago include cores, small sharp flakes for cutting and hammer stones for producing flakes and for pounding. Overwhelming evidence shows that these stone tools were used to cut up dead animals. Cut marks and percussion marks from breaking open long bones to extract marrow occur on hundreds of fossils. In 2002, Eric Hoberg showed that human-specific tapeworms are most closely related to tapeworms that infest the carnivores with which early *Homo* competed for carcasses nearly 2 million years ago. There is much evidence that early hominins ate meat, but little to show that their tools

R. MCVAY/GETTY

Learning how to control fire for cooking might have helped early hominins develop bigger brains.

NATURE|Vol 459|25 June 2009

were used on plants, and none that there were cooking fires or digging sticks at that time.

Wrangham interprets various anatomical traits of humans — small mouths, small gapes, weak jaws, small teeth and a short digestive system — as adaptations to eating soft, high-density food with little fibre in it. Are these truly adaptations to cooked food?

In 1995, Leslie Aiello and Peter Wheeler suggested that increased meat-eating caused the increase in hominin brain size about 2 million years ago. They proposed that a meat-rich diet accessed with stone tools permitted guts to become smaller, freeing up energy for brain growth. Primates with smaller guts consistently have larger brains, and humans are the extreme case of small guts and large brains, they said. Their hypothesis is compelling.

Wrangham argues that lumps of raw meat are difficult to digest and protein intake must be balanced with carbohydrate intake. Carbohydrates can be obtained from fat, marrow, brains, fruits and nuts, and are abundant in foods happily eaten raw by baboons, chimps and warthogs — such as corms, rhizomes and tubers of savannah plants — so Wrangham's dismissal of these sources of carbohydrates is unpersuasive.

Fire has other advantages, Wrangham observes. Sleeping on the ground at night was dangerous for hominins and a fire kept predators away. True, but antelopes sleep on the ground and have not yet invented fire or cooking.

Wrangham's analysis of the effect of cooking on humans is stronger. Cooking may have promoted a division of labour by sex. Large, prized food items hunted by males were taken home and shared, as were large vegetable items, such as tubers, gathered by females. Because females have babies, they often took on the tasks of foraging for staple vegetable foods and cooking.

Pair-bonding is another component of Wrangham's scenario. It ensures that the male hunter will have something to eat if he returns empty-handed, and reassures him that the baby he feeds is his. Pair-bonding assures the female gatherer that someone will contribute high-protein items to her baby. Because fires are visible for long distances, they may draw food thieves, so pair-bonding also gives protection to the cooks.

Wrangham's thinking about the effect of food choices on society is interesting, but his attempt to superimpose his hypothesis on to the early fossil and archaeological record is unconvincing. ■

Pat Shipman is professor of anthropology at Penn State University, University Park, Pennsylvania 16802, USA.
e-mail: pls10@psu.edu

Our changing body image

Assembling Bodies: Art, Science and Imagination
Museum of Archaeology and Anthropology, Cambridge, UK
Until November 2010

A gleaming transparent head, studded with stainless-steel screws and embedded with microchips, sits immovable on its plinth. Unnervingly, its left eye looks real: lidless, it flickers in mute appeal as if someone were trapped inside. *The Head of the Blue Chip II* by Dianne Harris is one of several artworks commissioned by the Museum of Archaeology and Anthropology in Cambridge, UK, to complement the historical — and indeed prehistoric — artefacts that make up its current exhibition, *Assembling Bodies*.

Harris's head speaks of a possible future in which the mind might become as readable as your computer's hard drive. Such a transformation would surely take place under academic scrutiny. But how can we know of the past changes in belief that have shaped our view of the human body, such as those that led our Bronze Age ancestors to cease burying dead bodies and instead cremate them?

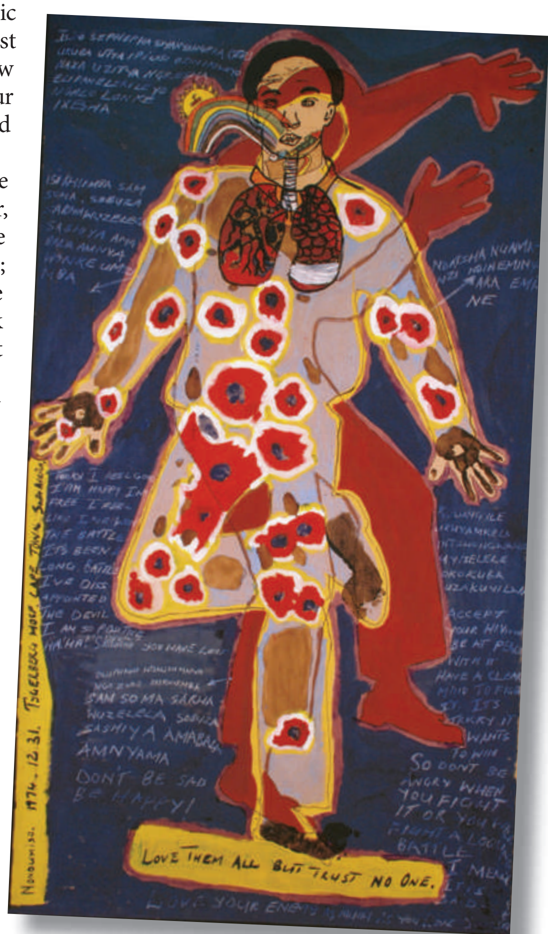
Scholars in different disciplines from the universities of Cambridge and Leicester, UK, have been funded by the Leverhulme Trust to explore these and other questions; *Assembling Bodies* is the publicly visible outcome. Curators Anita Herle, Mark Elliott and Rebecca Empson have brought ancient and modern cultural and scientific artefacts together with contemporary artists' responses to perennial questions about our material being. Clustered around seven themes, the exhibits, many loaned by other institutions, range from classical sculptures to a model of DNA, and from anatomical drawings and scientific instruments to an installation of string bags from Papua New Guinea.

The exhibition's section on genealogies contains one of the 119 volumes of the printed 'library' of the human genome commissioned by the Wellcome Trust. We see that attempts to codify human kinship in written form are not new, with a display of John Speed's early seventeenth-century Bible open at the genealogy he drew to link Adam to the Virgin Mary.

Measurement and classification were the foundations of anthropology as a professional discipline, and the museum draws on its own collections to show how the drive to define racial and psychological types could be subverted by its own methods. Alfred Cort Haddon's 1898 expedition to the Torres Strait islands between Australia and Papua New Guinea reported that, similar to modern geneticists, they found larger variations in physical measurements within the community of islanders than between the islanders and the expedition members. The sculptor Marguerite Milward made a collection of portrait busts in the 1930s to illustrate different Indian races, castes or tribes. She seems complicit with both colonial and Indian elites in justifying social stratification; yet standing in front of the pallid plastercasts, their individuality is what sticks in the mind.

What can and cannot be done with bodies is constrained by changing laws and customs. An Act for Preventing the Horrid Crime of Murder in Britain in 1751 used the threat of anatomical dissection as a deterrent to homicide.

MAA 2002.20, UNIV. CAMBRIDGE



South African artist Nondumiso portrays her experience of living with HIV in *Body-map* (2003).

It hangs on the wall across the room from a projection of almost 2,000 digitized cryo-sections that make up part of the US National Library of Medicine's Visible Human Project; they were taken from a Texan mechanic who received the death penalty for murder in 1993. The exhibition also includes a copy of the Guidance for the Care of Human Remains in Museums 2005. Appropriately, a funeral effigy from Vanuatu in the South Pacific, known as a rambaramp, contains the skull of its subject but is displayed with the permission of the community from which it was collected.

The idea that we all have multiple bodies, each made visible by the technologies used to explore them, is expressed in five life-sized body maps created by the Bambanani Women's Group in South Africa to document their experience of living with HIV/AIDS and its treatment. Within an outline tracing of her own body, each woman has illustrated her feelings about what is going on inside it, mixing biomedical, social, religious and moral perspectives.

Hanging above the central atrium is Jim Bond's *Anamorphic Man*. The parts of the sculpture, constructed of steel wire at vastly differing scales, appear disjointed and meaningless until the viewer moves to the single viewpoint from which the assembly coheres as a human figure. *Assembling Bodies* achieves the same trick, offering visitors an opportunity to make sense of the human enterprise of exploring and representing their material selves. ■

Georgina Ferry is a writer based in Oxford, UK, and author of *Max Perutz and the Secret of Life.*

Science journalism: From the newsroom

The Washington Post's national environmental reporter **Juliet Eilperin** and its executive editor **Marcus Brauchli** discuss the future of science coverage in their newspaper.

JE: Like many news organizations, declines in advertising dollars and profits have led to *The Washington Post* losing dozens of reporters and editors in the past five years through early retirement offers — four on the science beat alone. How is that changing things?

MB: The most obvious effect is that we're getting by with fewer people and being more selective in committing our resources. We're also creating a new health, science and environment unit headed by Frances Stead Sellers, who will oversee a team of reporters covering these subjects nationally and locally. This will eliminate the old divisions in our newsroom and provide stronger, better-coordinated coverage online and in the paper.

JE: How do you think science news stories have changed as a result?

MB: We're steering away from reporting just the facts towards analysis and contextual coverage that helps readers to understand crucial issues. We need to explain what's happening, but we also need to tell general readers and policy-makers what some event or discovery means. This is where we can bring a level of expertise to science coverage that most news organizations cannot.

The biggest challenge is to avoid bias. Yes, we encourage analysis and expertise from our writers; no, we don't want to use our interpretative skills and authority in a way that undermines our commitment to being open-minded and fair.

JE: For years, embargoed scientific papers have driven the news cycle. Should we be moving away from this sort of coverage to emphasize unique content instead?

MB: We'll continue to cover important scientific papers — and we'll always endeavour to put them in context, bringing the expertise of our science writers to the issues, rather than reporting on a single study's findings. But we also have to make choices: sometimes we'll run copy from the newswires rather than using our writers' time to cover something everybody else is covering. We've got a team of skilled reporters and we want to devote their energies to coverage that enhances the *Post*'s mission to be for and about Washington — to serve the people and policy-makers who live and work here.

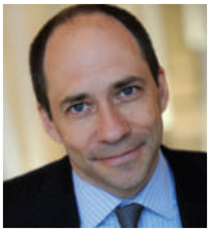
JE: That sounds as if there should be a heavy emphasis on political stories.

MB: Two of the biggest political stories this year are essentially scientific — climate change and health reform. Science is all about how we live, and you can't separate that from culture or politics.

JE: News organizations are devoting more resources to their web products — your online edition now has more readers than the print edition, and you're integrating the two staffs. Do you see a distinctive place for science coverage in that transition?

MB: The ability to reach and inform readers immediately through the web is both an advantage and a huge challenge for science writing. Science created computers, the Internet — things that move at great speed — and you might think of science as being well suited to that. But it can be incremental, too.

Look at the swine-flu story. Within hours, the blog we put together was packed with compelling coverage, fed by reporters from across the newsroom. The rapid pace of new developments — a flu case here, a school closed there — is well suited to the speed of the Internet, but then you have to guard against sensationalism, against making each advance seem more important than it is. At the same time there is a slower, incremental process of gathering information about what the virus is, or how dangerous it may become. This isn't as suitable for the web, but both trajectories need to be reported.


THE WASHINGTON POST


J. EWAN/THE WASHINGTON POST

Marcus Brauchli (top) and Juliet Eilperin.

JE: Do you think the web changes how stories are told?

MB: Good science coverage is about much more than writing a single story; it's about using graphics to explain difficult concepts and procedures, being ready to answer readers' questions online, and using video and photo galleries to enhance our traditional storytelling. Good science writing entertains readers even as it informs them. ■

Juliet Eilperin and **Marcus Brauchli** are at *The Washington Post*, Washington DC 20071, USA.

See Editorial, page 1033, and online at www.nature.com/news/specials/sciencejournalism.

NEWS & VIEWS

STRUCTURAL BIOLOGY

Protein dynamics from disorder

Joel R. Tolman

The functions of proteins are often crucially dependent on how they move, but measuring the absolute magnitudes of protein motions hasn't been possible. A spectroscopic method looks set to change all that.

In the half-century since our first glimpse of a protein structure at atomic resolution, structural biology has had a fundamental role in advancing our understanding of biological function at a molecular level. Yet although it is generally agreed that proteins must have some flexibility in order to perform their roles, uncertainty remains about how well static structural models capture the essential features underlying function. A picture may be worth a thousand words, but does it tell the entire story?

This is a surprisingly difficult question to answer unambiguously, because the picture of protein dynamics available from established experimental methods is incomplete. But Salmon *et al.*[1] report in *Angewandte Chemie* the absolute magnitude of motions spanning the entire submillisecond timescale for the peptide backbone of the protein ubiquitin. In effect, this work provides a quantitative measurement of how 'blurry' the picture of a protein is as a result of that protein's motion.

The requirement for protein flexibility was recognized early on by Max Perutz, who noted that there was no obvious route for ligand molecules to enter the active site of haemoglobin[2]. Since then, the need for protein flexibility has been demonstrated repeatedly by observations that some proteins adopt different structures when in complex with different ligands, or as a result of differences in the conditions used to crystallize them. Yet efforts to obtain a direct, quantitative description of protein dynamics in solution have generally suffered from limited spatial resolution or timescale sensitivity. It is now well established that all proteins undergo fast, picosecond-timescale motions of limited amplitude, but little is known about their motions on longer timescales (with the exception of a few specific cases on microsecond or slower timescales that involve large-scale motions)[3,4].

Nuclear magnetic resonance (NMR) spectroscopy is a powerful tool for studying the structure and dynamics of small proteins at atomic resolution. When this technique is used to study protein molecules that are weakly aligned in the magnetic field of the spectrometer, previously unobservable interactions, such as couplings between nuclear magnetic dipoles, manifest themselves as a splitting of the peaks in the spectrum. The splitting is proportional to the degree of alignment.

In the absence of dynamics, these so-called residual dipolar couplings (RDCs) between covalently bonded pairs of atoms depend on only two things: the extent of the molecule's overall alignment with the magnetic field, and the projection of the bond onto the direction of maximum molecular alignment with respect to that field (Fig. 1a). In the presence of dynamics, the movements of the bond occurring on a submillisecond timescale act as a source of disorder, diminishing the effective magnitude of alignment for that bond (Fig. 1b). But this second case is complicated because one cannot readily distinguish between the orientational and dynamic components of RDC measurements.

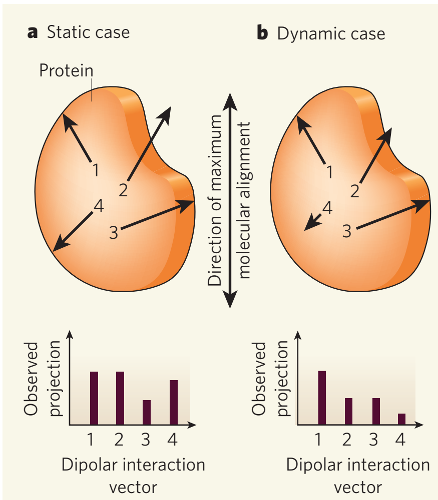


Figure 1 | Molecular alignment enables the detection of protein dynamics. Weak alignment of dissolved protein molecules in a magnetic field leads to the incomplete averaging of dipolar interactions between pairs of atoms, represented by the numbered vectors. These unaveraged interactions appear in NMR spectra as residual dipolar couplings (RDCs) — splitting of peaks in the spectrum that scale proportionally with the alignment strength, and that depend on the individual projections of dipolar interaction vectors onto the direction of maximum molecular alignment. **a**, In static molecules, RDC measurements of covalently bonded pairs of chemically identical nuclei depend only on the angle between the direction of maximum alignment and the specific bond. **b**, For dynamic molecules, submillisecond molecular motions introduce local disorder, which appears as an apparent reduction in alignment strength, as indicated by the shortened vectors. The resulting convolution of orientational and dynamic contributions to the RDC measurements can be untangled by performing several experiments in different solvents that cause the molecules to align in different ways. Salmon *et al.*[1] use this deconvolution approach in their method for analysing protein dynamics.

Fortunately, the orientational and dynamic components of RDCs can be separated by collecting several sets of data, with the protein aligned differently for each set[5,6]. But although this allows the relative magnitude of motions to be established, the absolute magnitude remains undetermined because of uncertainty in the overall magnitude of molecular alignment. Any underestimation of the magnitude of molecular alignment translates into an underestimation of the amplitude of internal motions.

Salmon *et al.*[1] report a method that overcomes this problem, and that allows the absolute amplitude of submillisecond motions of the peptide backbones of proteins to be established. Their approach uses RDC data from different bonds within each peptide unit, enabling the orientation and asymmetric motions of each individual peptide unit to be determined. Because of the asymmetry of the peptide motions, RDC data aren't compatible with any arbitrarily chosen magnitude of molecular alignment. The authors therefore determine the absolute amplitudes of peptide motions by using a cross-validation analysis in which the goodness of fit of the determined dynamic and orientational parameters with the overall magnitude of molecular alignment is gauged by its agreement with some fraction of the RDC data that has been excluded from the analysis.

The authors applied their method to ubiquitin, an unusually stable protein that lacks a catalytic function but binds to many other proteins. For much of the protein, the newly

NATURE|Vol 459|25 June 2009

determined motional amplitudes[1] are in general agreement with the known amplitudes of picosecond dynamics previously obtained by NMR[7]. The newly determined amplitudes also agree well with a molecular-dynamics simulation of the protein performed by the authors to validate their results.

More excitingly, using their technique, Salmon *et al.* observed motions that occur on a longer timescale (> 400 nanoseconds) than could be accessed by the molecular-dynamics simulation, specifically for a β-turn region and for a loop that lies close to ubiquitin's protein-binding interface. These motions could not have been observed using previously available NMR methods. The authors' results reinforce a picture of ubiquitin as a protein that has a rigidly packed, stable fold in which most of the flexibility is concentrated in regions responsible for its functional roles — protein recognition and assembly into covalently linked polyubiquitin chains[8–10] that serve as cellular signals. This defuses a long-standing controversy about whether globular, non-enzymatic proteins such as ubiquitin undergo any motions on timescales longer than nanoseconds.

Salmon and colleagues' method[1] opens the door to studies of other proteins that are expected to be more flexible than ubiquitin. At present, the main obstacle to using the authors' technique is that many different solutions of the protein must be prepared, each providing environments that cause the protein molecules to align differently. This can be quite challenging. Nevertheless, the method enables a straightforward average of the atomic coordinates of proteins to be recorded for movements occurring over the full submillisecond timescale, which clearly has important implications. Knowledge of the extent of protein movements could enable the determination of ensembles of structures that represent the conformational fluctuations of proteins, help to guide the design of ligands, and advance our understanding of protein catalysis and allostery (the regulation of protein activity by ligand binding to sites other than the functional binding site).

Dynamics is often used to explain apparent discrepancies between structural models of proteins and their known functions, yet there are few cases in which direct links between dynamics and function have been made. As experimental techniques for determining dynamics continue to improve, we can look forward to gaining deeper insight into the question of whether dynamics is more often essential or extraneous to protein function. ■

Joel R. Tolman is in the Department of Chemistry, Johns Hopkins University, 3400 North Charles Street, Baltimore, Maryland 21218, USA.
e-mail: tolman@jhu.edu

1. Salmon, L. *et al. Angew. Chem. Int. Edn* **48,** 4154–4157 (2009).
2. Perutz, M. F. & Mathews, F. S. *J. Mol. Biol.* **21,** 199–202 (1966).
3. Bustamante, C. *Annu. Rev. Biochem.* **77,** 45–50 (2008).
4. Mittermaier, A. & Kay, L. E. *Science* **312,** 224–228 (2006).
5. Meiler, J., Prompers, J. J., Peti, W., Griesinger, C. & Brüschweiler, R. *J. Am. Chem. Soc.* **123,** 6098–6107 (2001).
6. Tolman, J. R. *J. Am. Chem. Soc.* **124,** 12020–12030 (2002).
7. Tjandra, N., Feller, S. E., Pastor, R. W. & Bax, A. *J. Am. Chem. Soc.* **117,** 12562–12566 (1995).
8. Briggman, K. B. & Tolman, J. R. *J. Am. Chem. Soc.* **125,** 10164–10165 (2003).
9. Lange, O. F. *et al. Science* **320,** 1471–1475 (2008).
10. Peti, W., Meiler, J., Brüschweiler, R. & Griesinger, C. *J. Am. Chem. Soc.* **124,** 5822–5833 (2002).

GRANULAR MEDIA

Structures in sand streams

Detlef Lohse and Devaraj van der Meer

An ingenious experiment that involves dropping a costly, high-speed video camera from a height of several metres reveals how free-falling streams of granular matter, such as sand, break up into grain clusters.

It is common knowledge that when a stone is thrown into a pond, a jet of water shoots upwards. High-speed video imaging and even snapshots taken with short exposure times reveal that the jet breaks up into droplet patterns. Such patterns have inspired artists such as Andrew Davidhazy. What is both remarkable and intriguing is that very similar arrangements are observed in the streams that emerge when a stone is dropped into loose, fine sand[1]. The common view, which we have also adhered to, about the origin of this striking, liquid-like behaviour of streams of sand or other granular materials has been that the inelastic nature of collisions between the grains causes them to cluster together into larger structures[2–4]. But on page 1110 of this issue, Royer and colleagues[5] show that this view is wrong, and that forces hitherto believed to be too small to cause the clustering are at work.

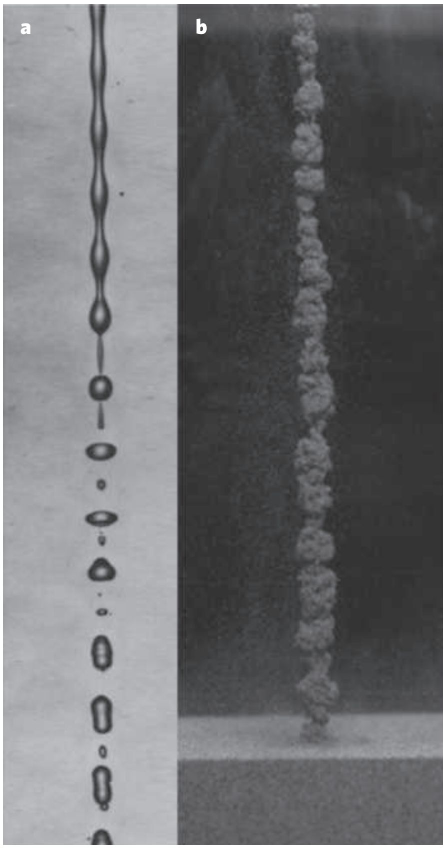
The formation of droplets in liquid jets (Fig. 1a) is caused by the Rayleigh–Plateau instability, which is driven by surface tension (the force that causes a liquid droplet to keep its shape). The understanding of this instability was pivotal to the conceptual development of hydrodynamics in the late nineteenth century by Joseph Plateau, Lord Rayleigh and their successors, and has since become textbook knowledge[6]. In contrast to liquid jets, granular streams of matter (Fig. 1b) were believed to lack surface tension. After all, granular matter such as sand is defined as a collection of grains that exert no forces on each other, with the exception of repulsive forces on collision[7]. It is thus surprising that streams of granular matter break up into structures that are similar to those of liquid streams.

Now Royer and colleagues[5] demonstrate that, in granular matter, tiny nanometre-range forces between the grains lead to a minute effective surface tension that, despite being 100,000 times smaller than that in water, can explain the clustering of grains in the jet. The researchers achieved this using an ingenious combination of nanometre-scale atomic-force-microscopy measurements of the forces between the sand grains with metre-scale tracking of the evolution of the granular streams. The tracking involved letting sand (and other granular material) stream out of a funnel and following its clustering dynamics, for almost a second, in the stream's co-moving frame of reference — that is, using a high-speed video camera that moved with the stream. This means that the researchers had to drop their US$80,000 high-speed camera from a height of several metres, a heart-stopping prospect — even though the impact of the camera on the ground was adequately buffered.

In their experiment, Royer *et al.*[5] observe a direct correlation between nanometre-range cohesive forces between the grains and the evolution of the jet structures within the first few metres of free fall. They find that modifying the nanometre-range forces — by varying the strength of intergrain cohesion through changes in the grains' surface roughness, the humidity or by using different materials — directly affects the break-up dynamics of the streams into clusters. In particular, the authors note that suppression of the nanometre-scale cohesion causes the clustering to vanish.

We would like to illustrate this striking phenomenon with two analogies. The first relates to the shapes of the grain clusters. As the authors[5] themselves note, the shapes of the clusters, including the double-cone necks at rupture, resemble the droplets that form in liquid nanojets and which were first found in molecular-dynamics studies that simulated injection of molecular fluids into a vacuum through a nanoscale nozzle[8]. Such studies show that the small number of molecules and thermal fluctuations in the nanojets lead to irregularities in the cluster shapes, very similar to those found in Royer and colleagues' granular jets, in which the number of sand grains is also small when compared with the huge number of molecules in macroscopic liquid jets.

The second parallel that can be made is with the emergence of planetary systems from dust grains in circumstellar gas disks. In this case, the first step is the formation of metre-sized boulders, but how this process continues to form kilometre-sized planetesimals is an



a, CAMBRIDGE UNIV. PRESS; b, AM. PHYS. SOC.

Figure 1 | Rupture of streams. a, In a process caused by surface tension, a water stream breaks up into droplets (image taken from ref. 10). **b**, Similarly, a stream of granular media such as a sand jet — which forms, for example, after the impact of an object on loose, fine sand — breaks up into clusters of grains (image taken from ref. 1). Royer and colleagues[5] show that, in granular media, the rupture and clustering process is driven by nanometre-range forces between the grains that lead to an effective surface tension that is 100,000 times weaker than in ordinary liquids.

unsolved problem[9], given that the (gravitational) attractive force between the boulders is so weak. However, Royer and colleagues' work teaches us that even minute forces can drive the formation of structures if the system is given enough time.

Traditionally, and in contrast to powders, one of the defining properties of granular matter has been the absence of cohesive forces between the grains[7]. But Royer and colleagues' observation of clustering in granular jets implies, conversely to what scientists have previously believed, that the distinction between powders and granular matter is much less clear-cut. Indeed, when velocity differences between the grains are small and timescales are considerable, even large grains can behave as powders.

The consequences of small, attractive forces may reach beyond the free-falling granular jets[1]. For example, the structures formed in the sand 'splashes' that are blown away after the impact of heavy objects on soft, fine sand[1] could presumably be correlated with the same type of cohesive force. Future studies are needed to investigate whether such a correlation exists, just as Royer and colleagues have done in this pioneering study of granular jets. ■

Detlef Lohse and Devaraj van der Meer are in the Physics of Fluids Group, Department of Science and Technology, J. M. Burgers Center, and at the Impact and Mesa+ Institutes, University of Twente, PO Box 217, 7500 AE Enschede, the Netherlands.
e-mail: d.lohse@utwente.nl

1. Lohse, D. *et al. Phys. Rev. Lett.* **93,** 198003 (2004).
2. Jaeger, H. M., Nagel, S. R. & Behringer, R. P. *Rev. Mod. Phys.* **68,** 1259–1273 (1996).
3. Kadanoff, L. *Rev. Mod. Phys.* **71,** 435–444 (1999).
4. Goldhirsch, I. *Annu. Rev. Fluid Mech.* **35,** 267–293 (2003).
5. Royer, J. R. *et al. Nature* **459,** 1110–1113 (2009).
6. de Gennes, P.-G., Brochard-Wyart, F. & Quere, D. *Capillarity and Wetting Phenomena: Drops, Bubbles, Pearls, Waves* (Springer, 2003).
7. Duran, J. *Sands, Powders, and Grains* 1st edn (Springer, 1999).
8. Moseler, M. & Landman, U. *Science* **289,** 1165–1169 (2000).
9. Johansen, A. *et al. Nature* **448,** 1022–1025 (2007).
10. Rutland, D. F. & Jameson, G. J. *J. Fluid Mech.* **46,** 267–271 (1971).

CELL BIOLOGY

A score for membrane fusion

Ruth N. Collins and Joshua Zimmerberg

Intracellular membrane fusion has been mimicked *in vitro* using a mix of 17 purified proteins and lipid bilayers. This technical tour de force allows the study of how cells orchestrate and perform such fusion events.

Molecules in living organisms are constantly being replaced, yet cellular structures can maintain their identity for a lifetime[1]. About half of all biological processes involve membrane proteins, which must be delivered, and eventually removed, with great accuracy to regulate the constancy of structural identity. This delivery and removal is mediated by the membrane-bound organelles of eukaryotic cells, which communicate with each other by budding off vesicles and other transport packages, which travel to, and fuse with, target membranes. The organization and regulation of membrane-fusion reactions, then, are crucial for virtually every membrane-bound biological process.

But how is fusion of internal membranes achieved? To understand these mechanisms, researchers have strived to devise *in vitro* systems that mimic fusion events *in vivo*. In this issue (page 1091), Ohya *et al.*[2] describe an enormous technical accomplishment — the self-assembly of 17 individual purified proteins into a fusion 'machine' whose activity and regulation recapitulate those of membrane fusion in an intact cell.

Two decades of investigation into membrane fusion have given us a set of key players. These include the SNARE family of membrane proteins and the Rab proteins, a subfamily of the Ras superfamily of GTPase enzymes. Mutations or pharmacological treatment of any of these components block fusion *in vivo*[3]. The membrane-bound SNARE proteins link the two membranes destined for fusion. They do this by interacting with partner SNARE proteins on the opposing membrane to form a stable, four-coiled bundle consisting of helices from several individual SNARE proteins[4] (Fig. 1, overleaf).

Rab GTPases, such as Rab5, are found both on membranes and in the cytosol. Rab proteins alternate between GDP-bound and GTP-bound forms, a switch controlled by proteins called guanine-nucleotide exchange factors. Like that of other members of the Ras superfamily, the nucleotide-dependent switch of Rab proteins is thought to control a downstream catalytic process. Rab proteins are involved in such processes as the regulation of organelle transport, the tethering of membranes before fusion, and the control of SNARE function[5].

Each of these protein families is served by a constellation of accessory factors; these include NSF, an ATPase enzyme that helps to recycle the SNARE protein complex; the Sec1/Munc18 (SM) proteins, which together with SNAREs are essential for fusion; and various effector proteins that interact with GTP-bound, active Rab[4,5].

SNARE complexes are structurally similar to viral envelope proteins known to catalyse membrane fusion[6], so a paradigm emerged casting these rod-shaped helical bundles into the central role of 'minimal fusion machine', with all other proteins assigned to supporting roles such as regulation of the SNARE complex[7]. However, direct tests comparing the contributions of SNAREs alone with SNAREs plus all the other proteins essential for fusion have not been performed because of the complexities of assembling microgram quantities of membrane proteins in defined lipid environments together with other fragile purified proteins — requirements that stretch the limits of current biochemical technologies.

Ohya *et al.*[2] overcome these obstacles, and reconstitute membrane fusion *in vitro* using physiological concentrations of 17 proteins derived from human membrane-bound organelles called endosomes, together with endosomal lipids. They quantify fusion by

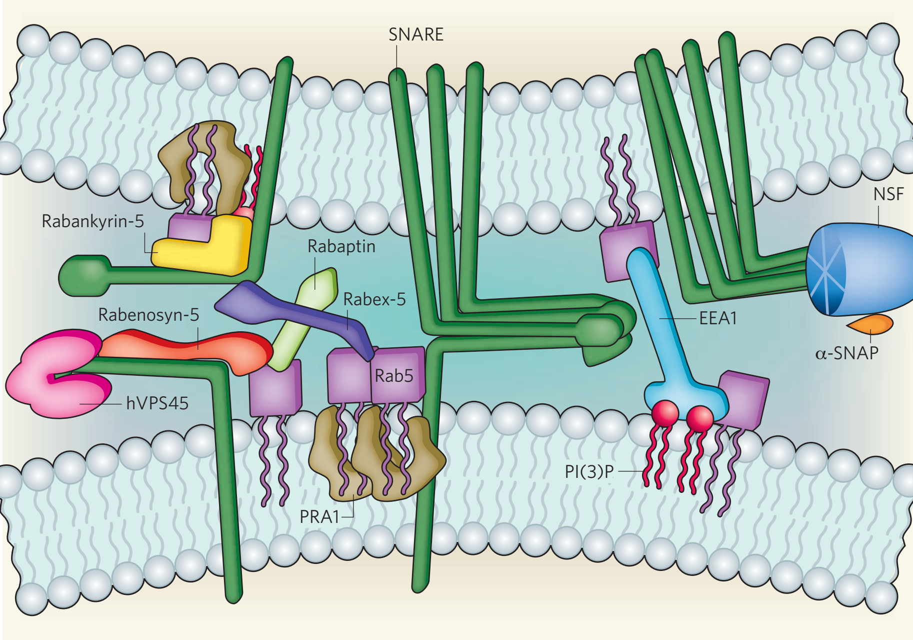


Figure 1 | Proteins work together to mediate endosomal membrane fusion. Ohya *et al.*[2] show that two protein systems, a Rab GTPase (Rab5) and the SNARE proteins, combine synergistically to drive endosomal fusion. Rab5 and SNAREs are assisted by multiple accessory factors. Rab5 effectors include rabankyrin-5, rabenosyn-5, rabaptin, rabex-5 and EEA1. SNARE accessory factors include α-SNAP, NSF and hVPS45. Not shown are the Rab–GDI complex (which retains Rab5 in the cytoplasm by masking its prenyl group), as this component is thought to function at an earlier stage, and another Rab5 effector, PI(3)K, which is instead represented by its lipid product PI(3)P. The complexity of the interactions observed in this supramolecular network is apparent; the recruitment of Rab5 effectors is stabilized by SNARE proteins and SNARE-protein accessory factors. For example, the Rab5 effector rabenosyn-5 also interacts with the SNARE accessory factor hVPS45. Linker proteins such as PRA1 bind to both Rabs and SNAREs.

using a novel content-mixing assay in which light is emitted when the contents of the fusing endosomes interact. This ensures direct detection of complete membrane fusion, unlike lipid-mixing assays that confuse fusion with hemifusion — fusion of only one monolayer of each of the two lipid bilayer membranes. The result is that the reconstituted fusion is comparable in both extent and rate to the fusion of intact biological endosomes purified from cell extracts.

The reconstituted fusion is robust enough for Ohya and colleagues to increase the diameter of the resultant vesicles by up to tenfold over their initial size (the equivalent of a 100-fold increase in surface area, or fusion of 100 vesicles into one fusosome). Fusion using this new system therefore differs from that observed in reconstitutions involving only SNARE proteins, in which small changes in the diameter of fusion products are hard won. Rab5 GTPase seems to cooperate with SNARE proteins to direct and drive the specificity and extent of membrane fusion, in addition to the traditional roles ascribed to Ras-related GTPases as master regulators[5]. The exclusion of Rab5 from the mix eliminates fusion, and the effect of Rab5 addition can be detected only when the other Rab5 effectors are also present. This *in vitro* experiment recapitulates genetic experiments[8] by showing the dependence of the fusion reaction on each participating protein, including the Rab proteins and their effectors. The study also reveals novel insights into the role of SNAREs and SNARE-protein accessory factors in stabilizing the recruitment of Rab5 effectors.

The view that emerges from Ohya and colleagues' work[2] is that membrane fusion at physiological rates involves the synergistic interactions of an ensemble of proteins, and is not simply the sum of the actions of individual proteins. One way in which interacting clusters of SNAREs, Rabs, Rab effectors and SM proteins may cooperate to effect fusion is by changing local membrane morphology and energetics. Membrane-associated proteins can combine with lipids to form fluid proteolipid domains that drive vesicle budding[9]. This new work[2] suggests that the endosomal-fusion machine may itself function as such a specialized domain, and be composed of the following: membrane proteins that self-associate (SNAREs)[10]; linker proteins such as PRA1/HuYIP3 that bind to both Rabs and SNAREs[11]; proteins such as Rab5 that are prenylated — a modification that favours clustering with unsaturated lipids in the membrane; and specific lipids with defined head groups, such as PI(3)P. It may help to have many different proteins all focusing their lipid-interaction energies on a hydrophobic patch of membrane to drive hemifusion[12], much as rings of proteins can drive hemifission[13]. Perhaps one of the recruited proteins inserts amphipathic peptides — those with both hydrophobic and hydrophilic regions — into the lipid bilayer. This could buckle the membrane, and lower the energy needed for fusion. The calcium sensor synaptotagmin is thought to promote SNARE-mediated fusion of synaptic vesicles in neurons by a similar mechanism[14].

Organelle fusion requires two nucleotide cycles: the GTP hydrolysis of the Rab cycle, and ATP hydrolysis to reset SNARE conformations after fusion is complete. We now have an excellent system in which to elucidate the functions of these nucleotide cycles in membrane fusion, their stoichiometries, and how they are thermodynamically coupled to fusion. We can also investigate how the reactions discriminate between different donor and acceptor membranes, and how fusion is linked to the translocation of Rabs between membranes and the cytosol. Solving these challenging problems will require a combination of approaches from fearless young scientists who are as well versed in the energetics and chemistry of membrane lipids as they are in the reconstitution of membrane proteins.

This study[2] changes our view of the defined and separate roles of SNAREs and Rab proteins in membrane fusion, with Rab GTPases as conductors and SNARE proteins as responsive musicians. Rather, it seems that the players collaborate with each other to coordinate their activities, as in a chamber orchestra. The properties of fusion may arise from complex systems interactions among many components (Fig. 1). The stage is now set to explore how these collaborations take place in cellular membranes, how the score is adapted for different ensembles, and how fusion and fission are coordinated with sorting and with the dispatch of components to specific destinations. ■

Ruth N. Collins is in the Department of Molecular Medicine, Cornell University, Ithaca, New York 14853, USA. Joshua Zimmerberg is in the Program in Physical Biology, Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, Maryland 20892-1855, USA.
e-mails: rnc8@cornell.edu; zimmerbj@mail.nih.gov

1. Schoenheimer, R. *The Dynamic State of Body Constituents* (Harvard Univ. Press, 1942).
2. Ohya, T. *et al. Nature* **459,** 1091–1097 (2009).
3. Ungar, D. & Hughson, F. M. *Annu. Rev. Cell Dev. Biol.* **19,** 493–517 (2003).
4. Wickner, W. & Schekman, R. *Nature Struct. Mol. Biol.* **15,** 658–664 (2008).
5. Grosshans, B. L., Ortiz, D. & Novick, P. *Proc. Natl Acad. Sci. USA* **103,** 11821–11827 (2006).
6. Skehel, J. J. & Wiley, D. C. *Cell* **95,** 871–874 (1998).
7. Weber, T. *et al. Cell* **92,** 759–772 (1998).
8. Singer-Krüger, B. *et al. J. Cell Biol.* **125,** 283–298 (1994).
9. Shnyrova, A. V. *et al. J. Cell Biol.* **179,** 627–633 (2007).
10. Zwilling, D. *et al. EMBO J.* **26,** 9–18 (2007).
11. Martincic, I., Peralta, M. E. & Ngsee, J. K. *J. Biol. Chem.* **272,** 26991–26998 (1997).
12. Kuzmin, P. I. *et al. Proc. Natl Acad. Sci. USA* **98,** 7235–7240 (2001).
13. Bashkirov, P. V. *et al. Cell* **135,** 1276–1286 (2008).
14. Martens, S., Kozlov, M. M. & McMahon, H. T. *Science* **316,** 1205–1208 (2007).

PLANETARY SCIENCE

Enceladus with a grain of salt

John Spencer

The observation that water plumes erupt from cracks on Saturn's moon Enceladus has fired speculation about a possible subsurface ocean. The latest searches for sodium salts point to the existence of such an ocean.

Do the spectacular plumes of water vapour and ice particles seen on Saturn's icy moon Enceladus come from liquid water just below its frigid surface? That is the fascinating question addressed by Postberg *et al.*[1] (page 1098 of this issue) using data from the Saturn-orbiting Cassini spacecraft, and by Schneider *et al.*[2] (page 1102) using ground-based telescopes. Despite their very different techniques, the two teams use the same key element, sodium, in their search for Enceladan water.

Of all the icy moons orbiting the giant planets of our Solar System, Enceladus, just 500 kilometres in diameter, is the only one (so far) where we can watch, in detail, geological processes as they happen. Four relatively warm 120-km-long fractures, informally called tiger stripes, slice across Enceladus's south pole and eject supersonic plumes of water vapour and ice particles thousands of kilometres into space[3–7] (Fig. 1). The ice particles populate a dust ring around Saturn — the E ring, which is much larger and fainter than the planet's better-known main rings — and the vapour populates a torus of atoms and molecules encircling the planet. The chemistry of the plumes is of intense interest not only because it provides a unique opportunity to sample the interior of an icy moon directly, but also because the interior of this particular moon provides a potential habitat for extraterrestrial life.

NASA/JPL-CALTECH

Figure 1 | Streaming into space. Saturn's icy moon Enceladus — backlit by the Sun and accompanied by a background star (top left) in this image taken by the Cassini spacecraft on 24 March 2006 — spews ice grains into space from jets at its south pole. The bright background sky, which makes the night side of Enceladus seem dark by contrast, is a small portion of Saturn's E ring, which is composed of ice grains accumulated from decades of output from the Enceladus jets. Analysis of these grains[1], and the gas that accompanies them[2], is providing new clues about a possible subsurface ocean that might supply the jets.

Life requires at least three ingredients: a source of energy, which in Enceladus is provided — at least in part — by tidal heating caused by the varying gravitational pull of its parent planet as Enceladus travels around its slightly eccentric orbit; a suitable mix of chemical elements, which seem to be present based on Cassini's analysis of the plume gases[4]; and liquid water. So the question of whether Enceladus's internal heat can provide that water, by melting a portion of the ice shell that comprises much of the moon's bulk, is one of the hot issues in planetary science today.

Sodium is a valuable tracer of possible liquid water for two reasons. First, it is highly soluble in water, and so any Enceladan water that has had prolonged contact with the moon's silicate core should be rich in sodium salts, like Earth's oceans. Second, when dispersed in its atomic form, sodium scatters sunlight with extreme efficiency at its resonant wavelength of 589 nanometres (the familiar orange–yellow colour of sodium street lights), and is thus easy to detect even in minute quantities.

Schneider *et al.*[2] use spectrographs on ground-based telescopes to search for sodium emission in Enceladus's gas plumes and Saturn's neutral torus. They find no sodium there to high precision, in striking contrast to the bright sodium emission seen in the output from Jupiter's volcanic moon Io, and even in the ultra-thin atmospheres of comets, Mercury and our own Moon. If there is salty water in Enceladus, some process must be very efficient in preventing most of the sodium from escaping into space.

Postberg *et al.*[1] focus instead on the ice grains from the plumes, using the Cosmic Dust Analyzer instrument aboard Cassini to determine their chemical composition directly as Cassini flies through the E ring. They find that, although all the grains are dominated by water ice, about 6% of them are quite salty, containing roughly 1.5% of a mixture of sodium chloride, sodium carbonate and sodium bicarbonate. The authors note that the grain

50 & 100 YEARS AGO

50 YEARS AGO

In 1956, as an experiment, an agency was started in the London area to put women graduates in touch with any suitable part-time work ... Many potential employers are prejudiced against part-time workers; it was felt that they might have more confidence if they could interview not one but a number of suitable candidates for each vacancy ... It also seemed clear that much of this work could be done at home and that it would be greatly welcomed, especially by mothers of young children ... Many employers were sympathetic to the idea and more jobs gradually became available ... Fewer permanent than temporary jobs are filled. This seems to reflect a general reluctance on the part of married women to commit themselves to permanent work. Women who have not held a job since marriage or since having children are sometimes uncertain how it will work out.

From *Nature* 27 June 1959.

100 YEARS AGO

Some very remarkable observations have been made from time to time during the last twenty years on the effect of chemical stimuli in bringing about abnormalities in developing embryos. The "Lithium larvae" of the sea-urchin and of the frog ... are familiar examples of this class of phenomena, but perhaps the most remarkable is the "Magnesium embryo" of the fish, *Fundulus heteroclitus* ... A large percentage of the embryos of this fish, when subjected during their development to the influence of magnesium salts dissolved in sea-water, are found to possess a single median or "cyclopean" eye in place of the ordinary pair. These embryos may hatch and swim about in a perfectly normal manner, and the single eye is evidently fully functional ... [The] results seem to indicate that the monstrous Cyclops of man and other mammals may not be due to germinal variation, but to some effect of environment during development.

From *Nature* 24 June 1909.

composition is similar to the expected composition of an Enceladan ocean[8], and therefore suggest that these particles are derived from direct freezing of salty water from that ocean at the plume source. They propose that the rest of the ice grains, which have very little salt, may come from direct condensation of the plume vapour.

So the question that naturally arises is: are Schneider and colleagues' no-sodium-emission observations in conflict with Postberg and colleagues' findings? The short answer is no. Although the ice grains in the E ring eventually vaporize and liberate their contents into the neutral torus, their total cargo of sodium, when diluted by all the sodium-poor grains and the sodium-free gas, would be too small to be detected by Schneider and colleagues' ground-based observations.

The simplest, and perhaps most likely, inference from Postberg and colleagues' particle-composition data is that the plumes are directly derived from salty liquid water. But it's always possible that we are being fooled: for instance, the plume sources might currently be too cold for liquid water to exist, and the plume might be excavating the sodium-rich ice grains from some long-frozen salt pockets in Enceladus's crust. We also can't be sure that, if there is water at the plume source, it is connected to a salty ocean — Schneider and colleagues point out that the water might originally be salt-poor, only becoming salty by preferential evaporation of the more-volatile water vapour. But in any case, those salty grains provide our current best smoking (or steaming) gun pointing to present-day liquid water near the surface of Enceladus.

The salty water cannot, however, be boiling explosively straight into the vacuum of space, otherwise the sodium would be carried along and would have been easily detectable by Schneider and colleagues. Instead, the plume production process must leave most of the sodium behind. Distillation of fresh water vapour from salty water happens over Earth's oceans, of course, and could happen on Enceladus too if the evaporation proceeded slowly in pressurized chambers, for example, rather than in a vacuum. The plumes might then be supplied by leakage from the chambers along narrow fractures leading to the surface. Postberg *et al.*[1] point out that water evaporation must occur slowly for another reason: to prevent the water from freezing due to rapid loss of latent heat during the evaporation process (see online Supplementary Information to ref. 1).

Cassini will make four more close fly-bys of Enceladus before mid-2010, with the chance of an additional twelve fly-bys up to 2015, if NASA approves the 7-year extended mission currently in the planning stages. Many more discoveries are therefore likely, and with each one Enceladus becomes more exotic. Our picture of its subsurface must now be expanded to include the possibility of misty ice caverns floored with pools and channels of salty water, lurking beneath the tiger stripes. What else may lurk in those salty pools, if they exist, remains to be seen.

John Spencer is at the Southwest Research Institute, Boulder, Colorado 80302, USA.
e-mail: spencer@boulder.swri.edu

1. Postberg, F. *et al. Nature* **459,** 1098–1101 (2009).
2. Schneider, N. M. *et al. Nature* **459,** 1102–1104 (2009).
3. Esposito, L. W. *et al. Nature* **456,** 477–479 (2008).
4. Waite, J. H. *et al. Science* **311,** 1419–1422 (2006).
5. Abramov, O. & Spencer, J. R. *Icarus* **199,** 189–196 (2009).
6. Spitale, J. N. & Porco, C. C. *Nature* **449,** 695–697 (2007).
7. Spahn, F. *et al. Science* **311,** 1416–1418 (2006).
8. Zolotov, M. Y. *Geophys. Res. Lett.* **34,** doi:10.1029/2007GL031234 (2007).

STEM CELLS

The stress of forming blood cells

Luc Pardanaud and Anne Eichmann

The first heartbeat is an important moment in an embryo's life. The biomechanical forces created by pulsatile flow promote the formation of haematopoietic stem cells that equip the body with its mature blood cells.

Haematopoietic stem cells (HSCs) are rare cells that can self-renew and generate all types of mature blood cell. Adult HSCs reside in the bone marrow and function throughout life, but during embryonic development they are first formed in close association with the endothelial lining of embryonic blood vessels, before the marrow appears. The first 'definitive' HSCs, which are able to form all of the blood-cell types if transferred into a lethally irradiated host (whose marrow is destroyed), are found in the ventral wall of the aorta, where they seem to bud from the endothelium into the aortic lumen (Fig. 1). HSCs are thus formed in close contact with flowing blood, and two groups, Adamo *et al.*[1] writing in this issue of *Nature* (page 1131) and North *et al.*[2] writing in *Cell*, now show that blood flow provides a biomechanical trigger for HSC formation.

Blood flow is initiated in the embryo when the heart starts beating: at this stage, blood consists mainly of plasma and primitive red blood cells. Flow is needed to deliver oxygen and nutrients to tissues, but it also generates biomechanical forces, including fluid shear stress, which is required for the correct development of the heart[3] and blood vessels[4]. The close proximity of developing HSCs to embryonic arteries, especially the aorta, suggested that blood flow might influence HSC formation, but this idea had never been directly tested.

Adamo *et al.*[1] examined the role of fluid shear stress on haematopoiesis in cultures of mouse embryonic stem (ES) cells. They show that shear stress increases the expression of Runx1, which activates the transcription of genes required for the development of functional HSCs[5]. Also, shear stress increases the potential of cultured ES cells to give rise to colonies of haematopoietic cells — a potential that is characteristic of haematopoietic progenitors and HSCs. Notably, only a shear stress similar to that estimated to occur in the embryonic aorta could induce increased Runx1 expression and colony formation in cell-culture experiments, indicating that the induction of Runx1 requires flow conditions similar to those observed in living embryos.

To evaluate the effect of blood flow on HSC formation *in vivo*, the authors[1] studied mice in which the *Ncx1* gene had been deleted. Mice with this deletion have a defect in a protein found in large amounts in heart cells that pumps sodium and calcium ions across the cell membrane. In such mice, heartbeat and circulation fail to develop[6]. Although the mutation is lethal, *Ncx1*-knockout mouse embryos live long enough to allow isolation of the tissue surrounding the aorta.

Examination of this mutant tissue showed that the expression of endothelial markers was similar to that of normal littermates, indicating that the mutant tissue had developed vascular endothelium. However, *Ncx1*-mutant tissue expressed decreased amounts of Runx1, and its ability to form haematopoietic colonies was markedly reduced. These defects could be overcome by subjecting the mutant cells to shear stress. Adamo and colleagues[1] conclude that the mutant tissue can generate haematopoietic precursors, but that their embryonic development had not been triggered because of lack of exposure to the appropriate mechanical forces.

The formation of HSCs in the aortic wall has been observed in all vertebrate species examined, including fish, birds, mice, pigs and humans[7]. North *et al.*[2] show that blood flow also initiates HSC development in zebrafish, providing further evidence that biomechanical forces are an evolutionarily conserved trigger for HSC formation. They used a chemical genetic screen to identify compounds that affect HSC formation. Zebrafish embryos were exposed to a panel of chemicals, and the expression of *runx1* and another HSC-specific transcription factor, c-*myb*, was measured. Compounds that induced dilation of blood vessels and increased blood flow raised *runx1* expression, whereas compounds causing

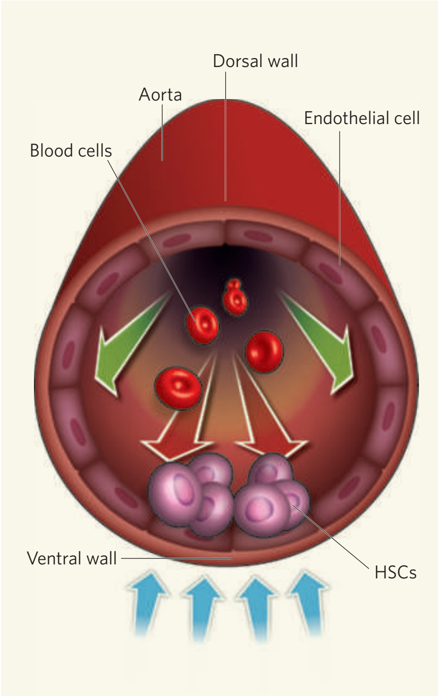


Figure 1 | Blood flow promotes development of haematopoietic stem cells (HSCs)[1,2]. HSCs and the aortic endothelium sense blood-flow-induced shear stress. HSCs bud into the aortic lumen from the endothelium lining the ventral part of the aorta. They develop only in the ventral part of the aorta, although shear stress is sensed throughout the aortic endothelium — ventrally (red arrows), laterally (green arrows) and dorsally. Signals from the mesoderm underlying the aorta (blue arrows) induce cells to produce HSCs and to respond to fluid shear stress.

constriction of blood vessels decreased its expression, indicating that blood flow can induce HSC formation. In agreement with this, *silent heart* zebrafish mutants, which lack a heartbeat and circulation, failed to form HSCs.

How does blood flow trigger HSC formation? Both studies[1,2] link blood flow to the nitric oxide (NO) signalling pathway, which is regulated by shear-stress-induced blood flow[8] and is a known modulator of haematopoiesis[9]. Inhibition of the enzyme nitric oxide synthase, with consequent reduction of NO signalling, decreased Runx1 expression and HSC formation in mice[1] and zebrafish[2], whereas addition of NO potentiated HSC formation in normal zebrafish and induced expression *of runx1* and c-*myb* in *silent heart* mutants[2]. North *et al.*[2] show that a specific nitric oxide synthase, Nos1, is responsible for HSC formation in zebrafish, whereas in mice, knockout of *Nos3*, the gene encoding endothelial Nos, leads to reduced formation of haematopoietic precursors.

Taken together, these studies[1,2] provide compelling evidence for an evolutionarily conserved, shear-stress- and NO-mediated pathway that leads to embryonic HSC formation. Pulsatile flow initiated by a regular heartbeat may induce NO production, thereby mediating HSC development in the embryonic aorta. An unresolved question concerns the relation between arterial specification of the endothelium — in which endothelial cells destined to line the aorta are induced to express specific genes — and the capacity to generate HSCs. Arterial specification is also induced by flow[10], and a possible alternative explanation for the results in both studies is that, in the absence of flow, arterial specification does not occur and HSC development fails because of the altered endothelial environment.

Moreover, only cells located on the ventral side of the aorta acquire an HSC fate; dorsally located cells, although exposed to the same shear stress, never give rise to HSCs (Fig. 1). Cells that make up the ventral and dorsal walls of the aorta have different developmental origins[11], and this may influence their competence to form HSCs or to respond to shear stress. In addition, cells at the ventral side of the aorta may be exposed to inductive signals from the underlying mesoderm (Fig. 1). Identification of the signals that allow ventral aortic cells to form HSCs in response to shear stress is thus a major challenge. Other outstanding questions concern the possible effects of the shear-stress response and NO function on HSC differentiation in the adult bone marrow, and whether manipulation of blood flow or NO signalling can open up new avenues for stem-cell therapy by promoting HSC formation after stem-cell transplantation. ■

Luc Pardanaud and Anne Eichmann are at Inserm U833, Collège de France, 75005 Paris, France.
e-mail: anne.eichmann@college-de-france.fr

1. Adamo, L. *et al. Nature* **459,** 1131–1134 (2009).
2. North, T. E. *et al. Cell* **137,** 736–748 (2009).
3. Hove, J. R. *et al. Nature* **421,** 172–177 (2003).
4. Lucitti, J. L. *et al. Development* **134,** 3317–3326 (2007).
5. North, T. *et al. Development* **126,** 2563–2575 (1999).
6. Koushik, S. V. *et al. FASEB J.* **15,** 1209–1211 (2001).
7. Cumano, A. & Godin, I. *Annu. Rev. Immunol.* **25,** 745–785 (2007).
8. García-Cardeña, G. *et al. Nature* **392,** 821–824 (1998).
9. Aicher, A. *et al. Nature Med.* **9,** 1370–1376 (2003).
10. le Noble, F. *et al. Development* **131,** 361–375 (2004).
11. Pardanaud, L. *et al. Development* **122,** 1363–1371 (1996).

NUCLEAR PHYSICS

Unexpected doubly magic nucleus

Robert V. F. Janssens

Nuclei with a 'magic' number of both protons and neutrons, dubbed doubly magic, are particularly stable. The oxygen isotope ^{24}O has been found to be one such nucleus — yet it lies just at the limit of stability.

Physicists often state that nuclear shell structure — the way in which protons and neutrons are arranged within a nucleus — is the cornerstone of any satisfactory description of an atomic nucleus. But over the past decade it has become apparent that the exact number of particles required to fill a particular shell is not as fixed as was once thought. The results of two experiments, one by Kanungo *et al.*[1] reported in *Physical Review Letters*, and the other by Hoffman *et al.*[2] in *Physics Letters B*, add significantly to the discussion. They demonstrate that ^{24}O, the oxygen isotope with proton number $Z = 8$ and neutron number $N = 16$, is a doubly magic nucleus. This result is all the more surprising because ^{24}O is also the heaviest oxygen isotope to exist.

The concept of nuclear shell structure is akin to that of atomic shell structure, in that shell closure results in enhanced stability. That is, there are certain nuclei — those that have a 'magic' number of protons and/or neutrons (2, 8, 20, 28, 50 and 82) — that have a full outer shell and are characterized by a large energy gap to the next available shell. As a result, they are more tightly bound than nuclei that have just one more proton or neutron. For neutrons, there is an additional such number: 126. Nuclei that have either the proton or the neutron number equal to one such magic number are thus termed magic nuclei; doubly magic nuclei are those with shell closures for both protons and neutrons.

The existence of nuclear magic numbers has been called into question as a result of studies of nuclei that — in terms of their proton and neutron numbers — are far from the region of stable, naturally occurring isotopes. It seems that the large energy gaps that give rise to the extra stability of magic nuclei are not all that robust and can change with proton and neutron number. In other words, there is experimental proof that some of these rather exotic nuclei, which are expected to be magic, are not particularly tightly bound, and other nuclei seem to signal the presence of new magic numbers[3].

The first indications of an unexpected shell closure at $N = 16$ — with the inference that ^{24}O might be a doubly magic nucleus — came from experiments investigating the binding energy of the neutron that can be most easily removed, and of the radioactive (β) decay properties of nuclei close to ^{24}O (refs 3, 4). However, the main properties associated with enhanced stability had thus far not been observed. Two things were lacking. First, a demonstration that the shell closure of ^{24}O has a spherical

shape, as expected for a tightly bound nucleus. And second, that ^{24}O is especially difficult to excite — that is, its first excited state is located at a high energy. This is where Kanungo *et al.*[1] and Hoffman *et al.*[2] check in with their experiments.

Both experiments used 'secondary' beams to probe the properties of ^{24}O. They started from a high-energy primary beam of the calcium isotope ^{48}Ca, which was made to interact with a beryllium target to produce a multitude of nuclear fragments. Fragment separators were then used to identify and select the species of interest, collect them into a beam — the secondary beam — and send them on to a reaction target. Kanungo *et al.*[1] studied the interaction of the very few (barely 3 particles per second) ^{24}O fragments produced in this way with a carbon reaction target. Specifically, they focused on the properties of ^{23}O products obtained from the direct removal of a neutron.

At the high energies involved, the ^{24}O secondary-beam particles undergo peripheral, grazing collisions with the target nuclei, where only the surfaces interact, and there is a simple relation between the physical descriptions, known as wavefunctions, of the incoming and outgoing nuclei[5]. From the momentum distribution of ^{23}O products, Kanungo and colleagues[1] were able to show unambiguously that the neutron removed from ^{24}O occupied, with a very large probability, the $2s_{1/2}$ energy level rather than the higher-energy $1d_{3/2}$ level (Fig. 1a). Because the wavefunction associated with the $2s_{1/2}$ state is spherical, this is an indication that the shell closure is a spherical one. Hence, the first of the two criteria for a doubly magic nucleus is fulfilled.

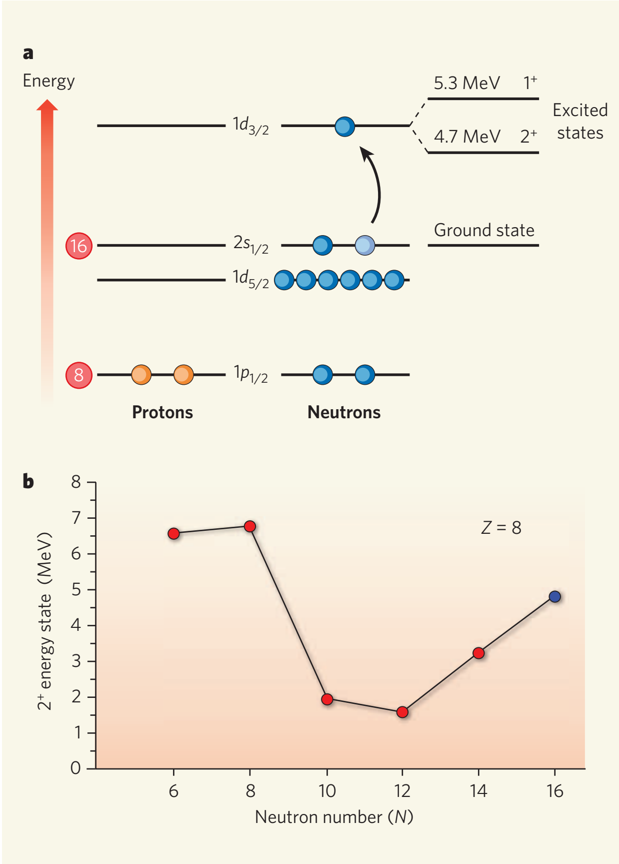


Figure 1 | Doubly magic nature of ^{24}O. Experiments by Kanungo *et al.*[1] and Hoffman *et al.*[2] provide evidence that ^{24}O (proton number $Z = 8$ and neutron number $N = 16$) is a doubly magic nucleus — that is, both its protons and neutrons are arranged into complete energy shells, and so have large energy gaps to the next available shell. **a,** In this representation of the nuclear shell structure of ^{24}O, only the last 2 of the 8 protons are shown: they occupy the $1p_{1/2}$ level, which is separated from higher-lying levels by a large gap, as 8 is indeed a magic number; similarly, the last 2 of the first 8 neutrons reside in the $1p_{1/2}$ neutron level, whereas the remaining 8 fill the $1d_{5/2}$ and $2s_{1/2}$ levels. Hoffman *et al.*[1] demonstrate that the first excited state of ^{24}O is a doublet of levels, termed 1^+ and 2^+, separated by a large energy gap from the $2s_{1/2}$ level. (In this notation, the subscripts 1/2, 3/2 and 5/2 correspond to energy-level multiplicities — or the maximum number of particles allowed per level — of 2, 4 and 6 respectively.) **b,** The variation of the energy of the 2^+ excited state with neutron number for different oxygen isotopes ($Z = 8$, $N = 6$–16) illustrates the sensitivity to shell closure: complete shells at $N = 8$ and $N = 16$ lead to an increase in energy. (Energies are given in megaelectronvolts (MeV); plot based on Fig. 4 of Hoffman *et al.*[2].)

Hoffman *et al.*[2] focused instead on the second criterion: the energy of the first excited state of ^{24}O. In their experiment, fluorine ^{26}F ($Z = 9$, $N = 17$) fragments were selected and turned into a secondary beam that was subsequently sent on to a beryllium target, removing a proton and a neutron from the fluorine fragments and thus leaving behind excited ^{24}O. The latter decayed promptly into ^{23}O and a neutron, for ^{24}O has no excited states bound to particle decay.

Hoffman and colleagues' detection[2] of both ^{23}O and a neutron represents a real tour de force: the ^{23}O ions had to be deflected from the direction of the secondary beam by a superconducting magnet, so that they could be detected without interfering with the forward-moving neutrons; these were measured using the modular neutron array (MoNA)[6] at Michigan State University. From approximately 400 ^{23}O–neutron coincidence events, Hoffman *et al.* showed that the neutron energy spectrum is best reproduced by simulations of the reaction that postulate the presence of a doublet of unbound excited states in ^{24}O, termed 1^+ and 2^+ states (Fig. 1a).

The authors[2] compared the energies of the first excited 2^+ states in all the $Z = 8$ oxygen isotopes that have an even neutron number N and showed how they vary markedly with N (Fig. 1b). As neutrons fill the $1d_{5/2}$ shell beyond the magic number $N = 8$, which corresponds to the doubly magic nucleus ^{16}O, the energy drops by a factor of about three on reaching $N = 12$ before increasing at $N = 14$, a manifestation that full $1d_{5/2}$ occupation is reached at that neutron number. A complete $2s_{1/2}$ shell at $N = 16$ leads to an even more dramatic increase in energy — a clear signature of the doubly magic character of ^{24}O.

As shown by Hoffman *et al.*[2], most theoretical shell-structure calculations are unable to reproduce the observations satisfactorily. The task is far from trivial, as several puzzling observations need to be reconciled. Experiments have demonstrated that oxygen isotopes such as ^{25}O or ^{26}O, which are heavier than ^{24}O, do not exist in nature — that is, they are unbound. Therefore, ^{24}O is truly remarkable because it is hard to excite, implying that it is doubly magic and very tightly bound. But it is located at the very limits of nuclear existence, as the addition of even a single neutron is not possible.

Equally surprising is the fact that the addition of a single proton, when moving from oxygen to fluorine, enables at least six additional neutrons to bind: observations indicate that even ^{31}F ($Z = 9$, $N = 22$) is bound[7]. Thus, the shell structure in Figure 1a changes drastically with proton and neutron number: it seems that, as soon as protons occupy the $1d_{5/2}$ orbital (as happens when going from O to F), the gap between the neutron $2s_{1/2}$ and $1d_{3/2}$ shells decreases significantly, an indication that a tensor force — an especially attractive, spin-dependent force between protons and neutrons — is providing the additional binding[8]. The full characterization of this force remains a challenge.

Experiments such as those of Kanungo *et al.*[1] and Hoffman *et al.*[2] highlight aspects of the physics of nuclei that are not readily apparent from the structure of stable nuclei, yet are essential for addressing the most fundamental challenge of nuclear physics — that of deciphering the exact nature of the nuclear force that binds protons and neutrons together in the nucleus and that defines the limits of the nuclear landscape. ■

Robert V. F. Janssens is in the Physics Division, Argonne National Laboratory, Argonne, Illinois 60439, USA.
e-mail: janssens@anl.gov

1. Kanungo, R. *et al. Phys. Rev. Lett.* **102,** 152501 (2009).
2. Hoffman, C. R. *et al. Phys. Lett. B* **672,** 17–21 (2009).
3. Warner, D. *Nature* **430,** 517–519 (2004).
4. Janssens, R. V. F. *Nature* **435,** 897–898 (2005).
5. Warner, D. *Nature* **425,** 570–571 (2003).
6. Baumann, T. *et al. Nucl. Instrum. Meth. A* **543,** 517–527 (2005).
7. Sakurai, H. *et al. Phys. Lett. B* **448,** 180–184 (1999).
8. Otsuka, T. *et al. Phys. Rev. Lett.* **87,** 082502 (2001).

Vol 459|25 June 2009|doi:10.1038/nature08122

Recent advances and emerging trends in plant hormone signalling

Aaron Santner[1] & Mark Estelle[2]

Plant growth and development is regulated by a structurally unrelated collection of small molecules called plant hormones. During the last 15 years the number of known plant hormones has grown from five to at least ten. Furthermore, many of the proteins involved in plant hormone signalling pathways have been identified, including receptors for many of the major hormones. Strikingly, the ubiquitin–proteasome pathway plays a central part in most hormone-signalling pathways. In addition, recent studies confirm that hormone signalling is integrated at several levels during plant growth and development.

Because plants have a sessile lifestyle, they must adjust to numerous external stimuli and coordinate their growth and development accordingly. The plant hormones, a group of structurally unrelated small molecules, are central to the integration of diverse environmental cues with a plant's genetic program. The 'classical' phytohormones, identified during the first half of the twentieth century, are auxin, abscisic acid, cytokinin, gibberellin and ethylene[1]. More recently, several additional compounds have been recognized as hormones, including brassinosteroids, jasmonate, salicylic acid, nitric oxide and strigolactones[2–7] (Table 1). Plants also use several peptide hormones to regulate various growth responses, but this class of hormones is beyond our scope here[8]. With the application of genetic approaches, mainly in *Arabidopsis thaliana*, many aspects of hormone biology have been elucidated. Most hormones are involved in many different processes throughout plant growth and development[1]. This complexity is reflected by the contributions of hormone synthesis, transport and signalling pathways, as well as by the diversity of interactions among hormones to control growth responses.

Genetic screens resulted in the identification of many of the proteins involved in hormone signalling and the analysis of these proteins has contributed significantly to our current models of hormone action. One particularly exciting outcome is the recent identification of receptors for auxin[9–11], gibberellin[12], jasmonate[13–15] and abscisic acid[16] (Fig. 1). Though far from complete, our improved understanding of hormone perception and signalling has allowed for comparisons between hormones. From these it is clear that some hormones (cytokinins, ethylene and the brassinosteroids) use well-characterized signalling mechanisms[17]. On the other hand, the identification and characterization of the auxin and jasmonate receptors, as well as proteins in gibberellin signalling, have highlighted a novel mechanism for hormone perception in which the ubiquitin–proteasome pathway has a central role[9–11,15,18].

In addition to these advances, the comparison of hormone signalling pathways between evolutionarily tractable members of the plant kingdom has yielded some important insights into the conservation and evolution of hormone signalling pathways. These comparisons have been facilitated by large-scale genome-sequencing projects such as those of *Physcomitrella patens* (moss), *Selaginella* (fern), *Arabidopsis thaliana* and *Oriza sativa* (rice). For example, the moss genome (an ancient plant ancestor) encodes proteins that function in auxin, abscisic acid and cytokinin signalling, whereas the genome of green algae does not, suggesting that these pathways emerged when plants were colonizing land[19,20]. In contrast, a comparison of the moss genome with more recently diverged plant genomes suggests that signalling mechanisms for gibberellin, ethylene and the brassinosteroids probably did not evolve until after the evolutionary split of moss and vascular plants[19,20]. These observations will be expanded as additional hormone signalling components are identified and more genome sequences become available.

This is an exciting time in the field of plant hormone biology because our knowledge of hormone biosynthesis, metabolism, transport, perception, signalling and response has grown exponentially over the past few years. As a result, recent reviews have been written for individual hormones covering topics from metabolism and transport to signalling[21–28]. Here, we review some of the advances in plant hormone signalling. We focus on newly identified hormone receptors and the broad role of regulated protein turnover in plant hormone

Table 1 | Plant hormone receptors

Hormone	Receptor type	Receptors	References
Auxin	F-box protein	TIR1, AFBs	9–11
Abscisic acid	G-protein, Chelatase	GTG1, GTG2, GCR2*, CHLH*	16, 55, 58
Cytokinin	Two-component regulators	CRE1, AHK2, AHK3	Reviewed in ref. 17
Gibberellins	Hormone-sensitive lipase like	GID1	12
Ethylene	Two-component regulators	ETR1, ERS1, ETR2, EIN4, ERS2	Reviewed in ref. 17
Brassinosteroids	Leucine-rich repeat receptor-like kinases	BRI1	Reviewed in ref. 17
Jasmonic acid	F-box protein	COI1	13–15
Salicylic acid	Unknown		
Nitric oxide	Unknown		
Strigolactones	Unknown		

[1]Molecular Kinetics Inc., 6201 La Pas Trail, Suite 160, Indianapolis, Indiana 46268, USA. [2]University of California San Diego, Section of Cell and Developmental Biology, 9500 Gilman Drive, La Jolla, California 92093, USA.

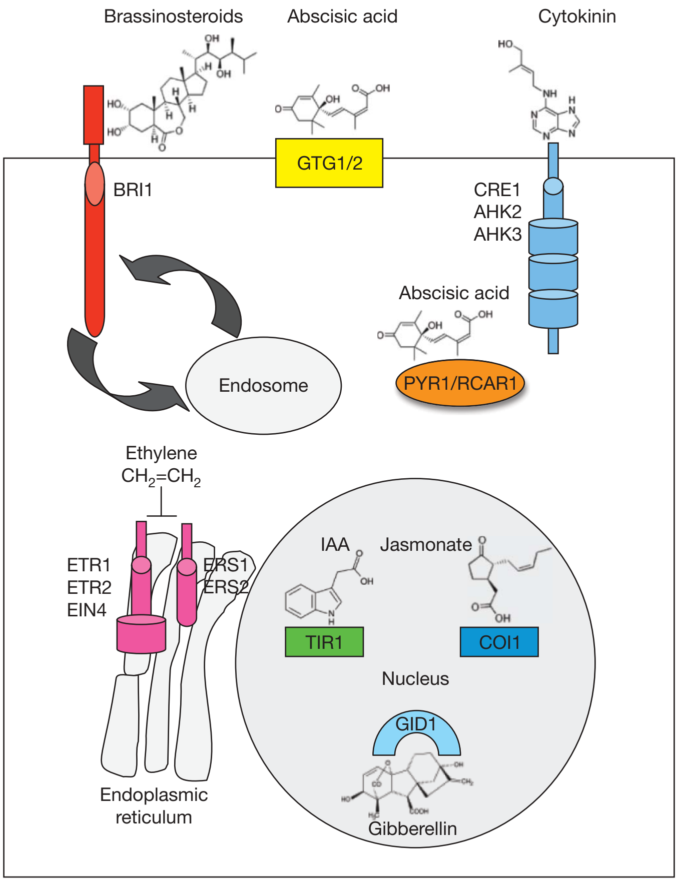


Figure 1 | Sites of plant hormone perception. BRI1 is a membrane-associated receptor that cycles between the plasma membrane and endosomal compartments. The extracellular leucine-rich repeat domain binds brassinosteroids and transduces the signal through an intracellular kinase domain. GTG1 and GTG2 are GPCR-type G proteins that bind abscisic acid. They have inherent GTPase activity but also interact with the only canonical Gα subunit in *Arabidopsis*. PYR1/RCAR1 is a soluble ABA receptor that represses PP2C phosphatases in the presence of ABA. The cytokinin receptors CRE1, AHK2 and AHK3 are plasma-membrane-associated and perceive cytokinin through their extracellular domains. Cytokinin binding triggers a phosphorylation cascade that is ultimately transmitted to response regulators in the nucleus. Like the cytokinin receptors, the known ethylene receptors are two-component regulators. All five receptors are active in the endoplasmic reticulum and transmit their signal through a common downstream component called CTR1. TIR1 and COI1 are F-box proteins that are integral components of SCF-type E3 ligases and recognize the plant hormones auxin and jasmonic acid respectively. GID1 is a nuclear-localized receptor for gibberellins. Gibberellin binding to GID1 results in the enhanced degradation of DELLA proteins.

signalling pathways. We also discuss some of the ways that hormone pathways are integrated during plant growth and development.

Auxin perception by a new class of receptor

Auxin is crucial in regulating plant growth and development from embryogenesis through maturity. As were most hormone signalling proteins identified in plants, the auxin receptors were first found through mutant screens. In this case the screen was for *Arabidopsis* seedlings with an altered response to auxin or auxin-transport inhibitors. Many of the auxin-resistant mutants identified in this way are disrupted in components of the Skp1/Cullin/F-box (SCF) ubiquitin ligases (E3) or in proteins that regulate SCF activity[29].

The E3 ligases are the last enzymes in the ubiquitin–protein-conjugation pathway and confer specificity to the pathway. In the case of SCF-type E3 ligases, the F-box protein interacts directly with the substrate and thus determines the substrate specificity of the complex[30]. SCFs were first implicated in auxin signalling with the identification of an F-box protein called TIR1. Recessive mutations in *TIR1* confer auxin resistance, implying that the protein is required for degradation of negative regulators of auxin response[31]. A key event in the characterization of the auxin-signalling pathway was the discovery that SCF^{TIR1} is directly linked to auxin-regulated transcription[32].

The auxin transcriptional response is controlled by two large families of transcription factors; the auxin/indole-3-acetic acid (Aux/IAA) proteins and the auxin response factors (ARFs) (of which *Arabidopsis* has 29 and 23 members respectively). ARFs bind the promoters of auxin-responsive genes and either activate or inhibit transcription depending on the type of ARF[33]. The Aux/IAA proteins bind to the ARFs through shared domains in both proteins called domains III and IV and repress auxin-regulated transcription[34]. Importantly, the Aux/IAA proteins are short-lived; their degradation is promoted by auxin and dependent upon TIR1. Many gain-of-function mutations in *Aux/IAA* genes have been isolated and in every case the mutations affect residues within a highly conserved region called domain II[34]. Biochemical studies demonstrated that domain II binds TIR1 and that this binding is enhanced by auxin[32,35,36].

Although these results suggested a mechanism for auxin-dependent de-repression of transcription, how auxin promotes the SCF^{TIR1}–Aux/IAA interaction remained unclear. Ultimately, TIR1 itself was shown to bind biologically active auxins directly and specifically[9,11]. Auxin binding to TIR1 increases the stability of the TIR1–Aux/IAA complex. Structural studies of TIR1 in the presence of auxin and a peptide encompassing domain II revealed how auxin promotes Aux/IAA degradation[37]. A single hydrophobic pocket on the surface of the leucine-rich repeat domain of TIR1 binds both auxin and the domain II peptide[37]. Auxin binds to residues at the base of this pocket and contributes to binding of the Aux/IAA protein[37]. Domain II of the canonical Aux/IAAs interacts with TIR1 residues directly above auxin, filling the remainder of the pocket[37]. One important implication of the structure is that both TIR1 and the Aux/IAAs appear to contribute to high-affinity binding of auxin. In this sense, it may be more appropriate to call TIR1 and the Aux/IAA protein co-receptors. If true, this also implies that different combinations of F-box protein and substrate may have unique auxin-binding characteristics.

Auxin research has a long history and the discovery that TIR1 functions as an auxin receptor was groundbreaking in several respects. The work indicates that F-box proteins, and perhaps other E3 ligases, can function as receptors for small molecules. Indeed, studies have demonstrated that this is probably true (see jasmonate signalling below). Further, the discovery that a small molecule can significantly enhance the interaction between an E3 and its substrate presents a new strategy for the development of drugs that target the ubiquitin–proteasome pathway[38]. Finally, detailed knowledge of auxin receptor function may stimulate the development of new plant growth regulators[39].

Recent results have also shed new light on the mechanism of Aux/IAA repression. Earlier studies showed that conserved domain I in the Aux/IAA proteins is required for transcriptional repression but the mechanism of repression was unclear[40]. Domain I of most Aux/IAAs contains an ethylene-response-factor-associated amphiphilic repression motif[40]. In 2008, a protein called TOPLESS (TPL) was shown to associate with domain I of the Aux/IAA protein IAA12 and to function as a transcriptional co-repressor[41]. These findings support an updated model in which the Aux/IAA proteins act as repressors of ARF-mediated transcription by recruiting TPL or related transcriptional co-repressors to the multi-protein complex[41]. Auxin de-represses transcription by promoting ubiquitination and subsequent degradation of Aux/IAA proteins through the action of SCF^{TIR1}. Without the Aux/IAA proteins, TPL is no longer associated with promoters of auxin-regulated genes (Fig. 2). Important questions about this model still remain. For example it is not known whether SCF^{TIR1} interacts with the Aux/IAA while in a complex with TPL and an ARF, with an ARF alone, or perhaps by itself. In addition, each of the relevant proteins is part of a large family (6 TIR1/AFBs, 29 Aux/IAAs, 23 ARFs, 5 TPL/TOPLESS

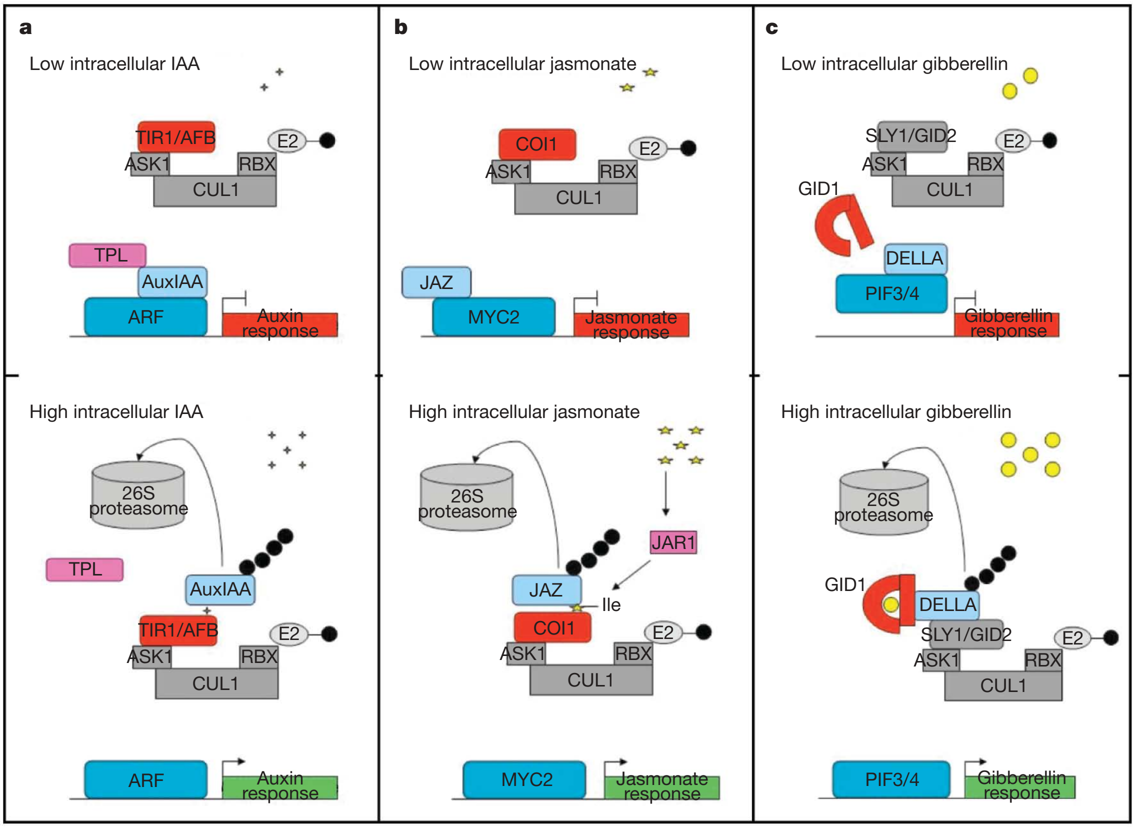


Figure 2 | SCFs are required for auxin, jasmonate and gibberellin signalling. a, The TIR1/AFB family of F-box proteins are auxin receptors. TIR1 is a component of the SCF complex that also consists of ASK, CUL and RBX. Auxin binding stabilizes the TIR1-AUX/IAA complex, resulting in degradation of the AUX/IAAs, which in turn releases TPL and permits ARF-dependent transcription. **b**, Binding of JA-Ile to COI1 promotes JAZ binding and ubiquitination. This results in de-repression of MYC2-dependent transcription of jasmonate-responsive genes. **c**, Gibberellin binding to the GID1 receptor promotes GID1–DELLA complex formation. GID1–DELLA binding promotes the interaction between the C terminus of the DELLA protein and $SCF^{SLY1/GID2}$. Degradation of the DELLAs promotes the release of PIFs, thus permitting DNA-binding gibberellin responses. In each panel, the hormone receptor is coloured red, the substrate protein is light blue, symbols representing hormones are yellow and components of the ubiquiting proteasome pathway are in grey. Black circles represent ubiquitin.

RELATEDs) and the potential specificity of interactions between different family members is just beginning to be explored.

Jasmonate perception is similar to that of auxin

It is now clear that TIR1 and its closest relatives the auxin-signalling F-box proteins (AFBs) serve as receptors for auxin[10]. There are roughly 700 F-box proteins encoded by the *Arabidopsis* genome. This raises the question: do any other F-box proteins function as receptors for plant hormones? Recent findings on jasmonate perception and signalling strongly suggest that jasmonate and auxin share a conserved mechanism of hormone sensing and response.

The oxylipin jasmonic acid and its metabolites, collectively known as jasmonates, are important plant signalling molecules that mediate biotic and abiotic stress responses as well as aspects of growth and development[23]. One of the mutants that helped define the role of jasmonate in plant growth is the *Arabidopsis coronatine-insensitive1* (*coi1*) mutant[42]. Coronatine is a phytotoxin that is structurally and biologically related to jasmonate. The *coi1* mutant is resistant to both coronatine and methyl jasmonate, and also confers male sterility[42]. Subsequent studies demonstrated that *coi1* mutants are perturbed in every aspect of jasmonate response, indicating that *COI1* has an essential role in jasmonate signal transduction. *COI1* encodes an F-box protein that is closely related to TIR1[43]. These results suggested that jasmonate response requires SCF^{COI1}-dependent degradation of repressors in much the same way as SCF^{TIR1} targets the Aux/IAAs. However, until 2007 the SCF^{COI1} substrates were unknown.

This changed when a novel family of transcriptional regulators called JAZ proteins (jasmonate ZIM-domain) was identified[13,14,44]. These groups showed that several full-length JAZ proteins are degraded in a proteasome-dependent manner following jasmonate treatment, but were stabilized in the *coi1-1* background, implicating SCF^{COI1} in the JAZ degradation pathway. Furthermore, members of the JAZ family were shown to interact with COI1 both *in vitro* and in a yeast two-hybrid test[13,14]. Indeed, subsequent biochemical analyses showed that radiolabelled coronatine binds to COI1-JAZ complexes with high affinity[15,45]. Interestingly, these studies also showed that jasmonate conjugated to isoleucine (JA-Ile) is the active molecule[45]. Taken together, the data suggest that SCF^{COI1} serves as a receptor for JA-Ile in order to stabilize the interaction between the F-box protein and its substrate.

The similarity between auxin signalling and jasmonate signalling does not end there. JAZ proteins do not have an obvious DNA-binding domain, suggesting that their effects on transcription could be indirect. Pull-down and yeast-two-hybrid assays indicate that the carboxy-terminal domain of JAZ3 (also known as JAI) interacts with sequences at the amino terminus of MYC2, a well-characterized transcription factor that modulates jasmonate-mediated transcription[13]. This raises the possibility that JAZ3/JAI directly blocks MYC2 function. JAZ3/JAI degradation via SCF^{COI1} in response to jasmonate would permit MYC2 to activate or repress downstream target genes in jasmonate signalling cascades. Several of the JAZ genes are themselves upregulated in response to jasmonate[13], indicating that a negative feedback mechanism may limit the response after jasmonate perception, again much like the auxin signalling pathway (Fig. 2).

Ubiquitination is a recurring theme

The newly identified receptors for auxin and jasmonic acid dramatically illustrate the importance of the ubiquitin–proteasome pathways in hormone signalling. However, it is also clear that targeted

protein turnover is an integral component in several other hormone signalling pathways.

Gibberellin signalling. Like auxin, gibberellins play a major role in diverse growth processes, including seed development, organ elongation and the control of flowering time[24]. Physiological data demonstrate that gibberellin responses are negatively regulated by DELLA proteins. (The DELLA proteins get their name from the conserved N-terminal DELLA domain[18].) Recent work exploring the integration of light and gibberellin signals during cell elongation has elucidated a model for DELLA-mediated growth regulation. DELLA proteins were shown to interact directly with the DNA-binding domain of two basic helix–loop–helix transcription factors (PIF3 and PIF4), sequestering them in inactive complexes[46,47]. Gibberellin accumulation destabilizes DELLAs, freeing PIF3 and PIF4 to activate the transcription of their target genes[46,47]. It is likely that DELLAs regulate the activity of several other transcription factors by this mechanism, as PIF3 and PIF4 are members of a larger subfamily of basic helix–loop–helix proteins that have similar DNA-binding domains.

Mutations within the DELLA domain of DELLA proteins result in a dominant gain-of-function gibberellin-insensitive phenotype[18]. Further, the DELLAs were found to accumulate in *Arabidopsis* and rice mutants defective in the F-box protein genes *sleepy1* (*sly1*) and *gibberellin-insensitive dwarf2* (*gid2*), respectively[48,49]. Thus, as observed in auxin and jasmonate signalling, gibberellin appears to regulate the abundance of a transcriptional repressor family by promoting ubiquitination through the activity of SCF-type ligases.

The gibberellin receptor was first identified in a genetic screen for signalling mutants in rice[12]. The *gibberellin-insensitive dwarf 1* (*gid1*) mutant is completely insensitive to gibberellin-treatment, suggesting a prominent role for the protein in gibberellin signalling[12]. The GID1 protein is nuclear-localized and binds biologically active gibberellins[12]. Three orthologous genes—*GID1a*, *GID1b* and *GID1c*—were identified in *Arabidopsis* and mutations in each gene were obtained and combined to study their contributions to the gibberellin response[50,51]. The *gid1a–c* triple mutant is gibberellin-insensitive, suggesting that all gibberellin responses require functional GID1 proteins in both rice and *Arabidopsis*[51].

The GID1 proteins interact with DELLA proteins in a gibberellin-dependent manner[12,51,52]. Furthermore, co-expression of the GID1 receptor enhances the interaction between DELLA proteins and the F-box proteins SLY1/GID2[51]. These data suggest that DELLAs are better able to interact with $SCF^{GID2/SLY}$ while in a complex with gibberellin-bound GID1. This interaction ultimately leads to ubiquitination and degradation of the DELLA repressor, thus promoting gibberellin-mediated transcription (Fig. 2).

The structures of rice GID1 and *Arabidopsis* GID1a have recently been reported and support this model[53,54]. GID1 has sequence and structural similarities to members of the hormone-sensitive lipase family. Like hormone-sensitive lipases, the GID1 primary structure forms a deep binding pocket whose access is controlled by an N-terminal flexible lid[53,54]. Gibberellin binding to GID1 probably induces the protein to adopt a compact form, with the N-terminal lid folding back over the gibberellin-binding pocket[53,54]. DELLA–GID1 interaction is achieved through three conserved motifs in the DELLA domain (DELLA, VHYNP and LExLE) that directly contact the N-terminal lid of GID1[53]. It is not yet clear how complex formation enhances the DELLA-$SCF^{SLY/GID2}$ interaction, but it may involve a conformational change in the C-terminal GRAS domain of the DELLA proteins[53].

Abscisic-acid signalling. Much recent work has elucidated putative mechanisms of abscisic-acid perception. A number of potential abscisic-acid receptors have recently been described but some of these studies have been controversial. One candidate, magnesium protoporphyrin-IX chelatase H subunit (CHLH), also known as GUN5 in *Arabidopsis*, was reported to have specific abscisic-acid-binding activity[55]. This group also presented data indicating that *Arabidopsis* plants with decreased CHLH levels have abscisic-acid-insensitive germination and stomatal aperture phenotypes, while plants overexpressing CHLH are hypersensitive to abscisic acid in these assays. However, GUN5 is an unconventional site of abscisic-acid perception because it is localized to chloroplasts[55,56]. Further, a recent study reported that barley CHLH did not bind abscisic acid and that mutants with reduced CHLH levels did not display an abscisic-acid phenotype, casting doubt on the idea that this protein is an abscisic-acid receptor[57].

The second candidate abscisic-acid receptor is a G-protein-coupled receptor called GCR2[58]. This suggestion is also controversial because the presence of a proposed transmembrane domain in GCR2 has been disputed[59]. Furthermore, genetic analysis of *gcr2* mutants failed to detect an abscisic-acid-related phenotype and some biochemical studies indicate that GCR2 does not bind abscisic acid after all[60–62]. At this point, it seems unlikely that GCR2 functions as an abscisic-acid receptor.

Three very recent papers show more promise. In a study from the Assman laboratory, another pair of G-protein-coupled receptors have been implicated in abscisic-acid response, GTG1 and GTG2[16]. Genetic evidence indicates that GTG1 and GTG2 act redundantly to mediate abscisic-acid responses during germination, flowering, root elongation and stomatal closure[16]. Perhaps more importantly, GTG1 and GTG2 were demonstrated to bind biologically active abscisic acid specifically in *in vitro* binding assays. Furthermore, the abscisic-acid binding dissociation constants of ~35 nM for GTG1 and ~41 nM for GTG2 fall within a physiologically relevant range.

Finally, two other papers present compelling evidence that a family of START proteins function as abscisic-acid receptors[63,64]. One of these studies began with the discovery of a selective abscisic-acid agonist called pyrabactin[63]. A genetic screen for pyrabactin targets led to the identification of the *PYRABACTIN RESISTANCE 1* (*PYR1*) gene, encoding a cyclase subfamily member in the START domain superfamily. Mutants deficient in *PYR1* and at least two additional members of the *PYR1-LIKE* (*PYL*) family are strongly abscisic-acid-resistant[63]. To learn more about the function of these proteins, the Cutler group[63] performed a yeast-2-hybrid screen using PYR1 as bait. They recovered HAB1, a member of a family of PP2C protein phosphatases known to regulate abscisic-acid signalling. The best characterized of these is ABI1. Strikingly, the interaction between PYR1 and the PP2C proteins depends on abscisic acid. Further, the PYR1 binding inhibits the phosphatase activity. At the same time, the Grill group were screening for ABI1 interacting proteins in yeast[64]. They identified a protein called RCAR1, which turned out to be identical to PYL9. Like PYR1, RCAR1 interacts with ABI1 in an abscisic-acid-dependent manner and inhibits its activity. Together, these studies suggest a fascinating model in which abscisic-acid signalling requires repression of PP2Cs through the action of PYR1/RCAR1 and related proteins[63,64]. At this point the relationship between the GTGs and PYR1/RCAR1 is unknown. However, abscisic-acid responses are diverse and it is possible that this diversity requires multiple sites of perception[65].

The intermediate signalling steps between abscisic-acid perception and response are unclear but the ubiquitin–proteasome pathway is known to be important (Fig. 3). Two RING E3 ligases, ABI3-Interacting Protein (AIP2) and Keep on Going (KEG), promote normal abscisic-acid signalling by regulating the abundance of abscisic-acid-responsive transcription factors, namely ABA-Insensitive 3 (ABI3) and ABA-Insensitive 5 (ABI5)[66,67]. ABI3 levels are reduced as abscisic-acid levels increase in cells. The detailed mechanism is unknown but there is evidence suggesting that abscisic acid increases AIP2 transcript and protein levels[66]. Increasing AIP2 levels then promotes the ubiquitination and degradation of ABI3, contributing to its accelerated degradation. Conversely, abscisic acid protects ABI5 from ubiquitin-mediated degradation, causing ABI5 levels to increase in response to the hormone. KEG has been implicated in the regulation of ABI5 and abscisic-acid signalling may prevent either

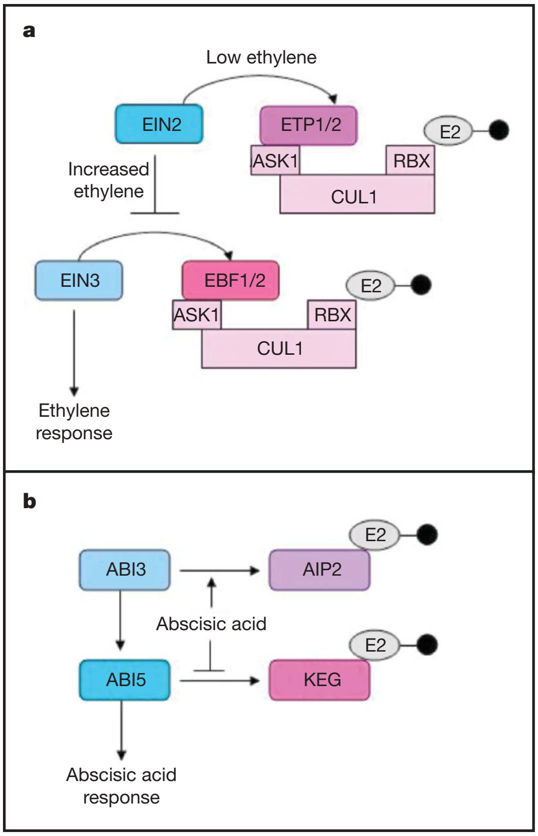


Figure 3 | E3 ligases in ethylene and abscisic-acid signalling. a, Two key proteins in the ethylene signalling pathway are regulated by SCF-type E3 ligases. In the absence of ethylene, EIN2 levels are kept low by SCF^{ETP1} and SCF^{ETP2}. As ethylene levels increase, *ETP1/2* expression is reduced and EIN2 levels increase. Concomitantly, EIN3 levels also increase, in part due to the effects of EIN2 on $SCF^{EBF1/2}$-mediated degradation of EIN3. **b**, ABI3 and ABI5 levels are controlled by the RING E3 ligases AIP2 and KEG, respectively. Abscisic acid promotes the ubiquitination of ABI3 and inhibits the ubiquitination of ABI5 to promote abscisic-acid responses. Black circles denote ubiquitin.

the recognition or the ubiquitination of ABI5 by KEG[67]. *keg* seedlings have a severe phenotype that includes growth arrest immediately following germination. Reducing ABI5 levels in *keg* mutants rescues some but not all aspects of the phenotype, suggesting a broader role for KEG in abscisic-acid signalling that may also extend to other abscisic-acid-responsive transcription factors[67].

Ethylene signalling. Ethylene signalling is also dependent on regulated protein turnover. ETHYLENE INSENSITIVE 3 (EIN3) is a positive transcriptional regulator required for transcription of *ETHYLENE RESPONSE FACTOR* genes[68]. EIN3 levels are regulated through the action of at least two related F-box proteins, EIN3-Binding F-box 1 (EBF1) and EBF2[69,70]. SCF^{EBF1} is thought to repress EIN3 levels when ethylene is low[71]. Two putative functions have been suggested for SCF^{EBF2}. SCF^{EBF2} may prevent excessive accumulation of EIN3 or perhaps remove it when ethylene levels decrease (Fig. 3). Normal ethylene responses are thought to decrease the degradation of EIN3 by SCF^{EBF1} and SCF^{EBF2}, thereby increasing EIN3 levels within responding cells[71].

More recently, another pair of F-box proteins—ETP1 and ETP2—have been shown to promote degradation of the ethylene signalling protein EIN2 in the absence of ethylene[72]. When ethylene is present, expression of *ETP1* and *ETP2* is reduced, allowing accumulation of EIN2. Thus the ubiquitin pathway is involved in ethylene signalling at multiple points in the pathway.

Strigolactone signalling. For many years auxin and cytokinin were thought to be the major mediators of shoot branching. However, studies of series of mutants including *ramosus* (*rms*) of pea, *dwarf* (*d*) of rice, *more axillary growth* (*max*) of *Arabidopsis*, and *decreased apical dominance* (*dad*) of petunia indicated that axillary bud outgrowth is also inhibited by an unidentified hormone[73]. Several affected genes were identified and found to encode CAROTENOID CLEAVAGE DIOXYGENASE 7 (*max3*, *rms5*, *d17*) or CCD8 (*max4*, *rms1*, *d10*, and *dad1*), suggesting that the hormone was related to carotenoids[73]. Very recently, two independent groups showed that levels of strigolactone, a carotenoid derivative, were reduced in rice *d* mutants and the pea *rms1* mutant[6,7]. Treating the respective rice and pea mutants with the synthetic strigolactone GR24 restored axillary bud outgrowth inhibition, suggesting that strigolactones are branching hormones[6,7]. In the context of shoot branching, it is interesting to note that an F-box protein called MAX2/RMS4 is required for response to strigolactone[74,75]. We expect to learn soon whether MAX2/RMS4 is the strigolactone receptor.

Integration of hormone signalling

Crosstalk between hormones is a very active area of research that has benefited from the recent elucidation of hormone signalling pathways. Although our knowledge of the molecular components and pathways that mediate hormone responses has improved enormously in recent years, the molecular mechanisms of hormone interaction remain poorly understood. The coming years should see a greater focus on the function of each hormone pathway in the context of larger regulatory networks.

Evidence for hormone crosstalk comes largely from the analysis of mutant phenotypes. Frequently mutants that are affected in one hormone pathway also display changes in other hormone responses. For example, auxin-related mutants such as *tir1*, *aux1* and *pin2* exhibit altered response to other hormones, including ethylene and abscisic acid[76,77]. From these studies and many others it is clear that hormone crosstalk is as complex as it is important. Almost certainly, all plant hormones interact with one or more additional hormones by affecting synthesis, transport or response. The type of interaction often depends on the tissue, developmental stage and environmental conditions. This presents a nearly endless number of possible opportunities for regulation. Even though the documented instances of crosstalk are numerous, our current knowledge of hormone interaction is quite sketchy. Here we will discuss some of the broad strategies for integrating hormone responses through examples from the recent literature.

It has become clear that diverse mechanisms have evolved to coordinate activity of hormonal pathways during development (Fig. 4). A prominent example is the regulation of hormone metabolism by an interacting hormone. This mechanism is exemplified by ethylene and auxin. Measurement of the rate of auxin biosynthesis after ethylene treatment revealed that ethylene stimulates the auxin biosynthetic pathway. Indeed, several genes required for auxin biosynthesis are under the transcriptional control of ethylene. These genes encode both the alpha and beta subunits of anthranilate synthase (*ASA1* and *ASB1*) and a newly identified family of tryptophan aminotransferases (TAA1)[78–80]. These enzymes function in tryptophan synthesis and the indole-3-pyruvic auxin biosynthetic pathway, respectively.

Auxin can influence ethylene biosynthesis as well. One of the rate-limiting enzymes in the ethylene biosynthesis pathway is 1-aminocyclopropane-1-carboxylate synthase (ACS)[81]. In *Arabidopsis*, there are nine ACS genes that can homodimerize or heterodimerize to form active enzymes[82]. Several of the ACS genes have been found to be regulated by auxin treatment at the transcriptional level[81]. Auxin has also been implicated in the production of jasmonic acid in flowers. Mutant analysis of *arf6 arf8* double mutant plants demonstrated a role in the coordination transition from immature to mature flowers[83]. During floral development, jasmonate levels peak just before anther dehiscence and then decrease[83]. Direct measurements showed that jasmonate production was reduced in *arf6 arf8*

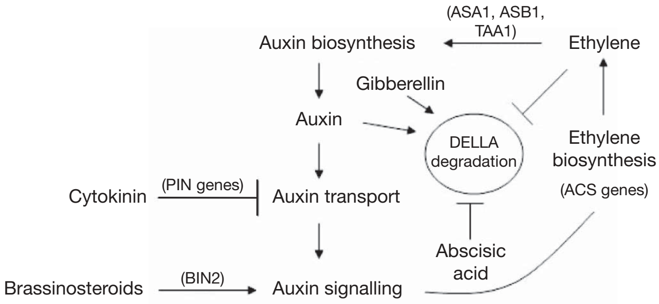


Figure 4 | Hormone integration. Auxin and ethylene maintain an important balance in cells, in part by regulating synthesis. Both hormones exert transcriptional control over critical genes in biosynthesis. A second mechanism of crosstalk is through the control of hormone transport. Cytokinin and auxin are antagonistic during lateral root initiation. To inhibit lateral root formation, cytokinin represses *PIN* expression, thereby inhibiting formation of an auxin gradient required for lateral root initiation. Brassinosteroids and auxin coordinate growth, in part by regulating an overlapping set of genes. In addition, phosphorylation of ARF2 by a brassinosteroid-regulated kinase modulates the activity of ARF2. Finally, auxin, ethylene and abscisic acid all converge on the DELLA proteins.

mutants at every stage of floral development[83]. Furthermore, gene expression studies revealed that several known jasmonate biosynthetic genes were underexpressed in *arf6 arf8* mutants, revealing a regulatory role for auxin in jasmonate production[83].

Hormone interactions also occur at the level of hormone distribution. An illustrative example of this type of crosstalk is the opposing action of auxin and cytokinin during lateral root initiation. It is widely known that polar auxin transport and the establishment of an auxin gradient is a very important determinant of plant growth and morphological patterning[84]. During root development, auxin promotes lateral root initiation while cytokinin opposes this response. One way that cytokinin exerts this effect is to influence the expression of the *PIN* auxin-efflux carrier genes[85]. At least five *PIN* genes work collectively to establish auxin gradients in roots by controlling the direction of polar-auxin transport[86]. By reducing *PIN* expression, cytokinin disrupts the local auxin gradient formation in lateral root founder cells, thereby inhibiting lateral root initiation[85].

Hormonal signalling pathways are also known to interact at the level of gene expression. For example, studies show that there is significant overlap between auxin- and brassinosteroid-responsive gene sets[87,88]. Generally, common target genes repressed by auxin are also repressed by brassinosteroids, and genes induced by auxin are induced by brassinosteroids, suggesting coordination between the signalling pathways. Furthermore, transcriptional profiling in the brassinosteroid-deficient mutant *brx* showed that very few auxin response genes responded normally to auxin[89]. Conversely, many brassinosteroid-responsive genes are mis-regulated in the *yucca* mutant that overproduces auxin[87]. Taken together, the data suggest that auxin and brassinosteroid signalling pathways often converge on a set of common target genes. A molecular mechanism for this convergence was recently elucidated in which the brassinosteroid-regulated BIN2 kinase directly regulates ARF2 activity[90]. ARF2 is an auxin-response factor that inhibits the transcription of auxin-responsive genes. The phosphorylation of ARF2 by BIN2 disrupts DNA binding, leading to inactivation of ARF2 and a subsequent increase in transcription of auxin-responsive genes[90].

Control over key components of signalling pathways by other hormone signals is another common example of cross-talk strategy. As described above, DELLA proteins are central regulators of the gibberellin-mediated signalling pathway and appear to be a common crosstalk node for several interacting hormones, including auxin, ethylene and abscisic acid[91]. Gibberellin signalling during root elongation is known to require auxin because disruption of polar auxin transport or signalling diminished the effects of gibberellin on root elongation[92]. The attenuated growth response corresponded with reduced RGA (a specific DELLA protein) degradation in root cells. These observations indicate that auxin promotes the gibberellin-induced destabilization of some of the DELLA proteins to affect gibberellin responses[92].

Similarly, ethylene may also target DELLA proteins to exert antagonistic actions with gibberellin during root growth[93]. The gibberellin-insensitive *gai/rga* mutant exhibits ethylene-insensitive root growth, indicating that ethylene regulates root growth in a DELLA-dependent manner[93]. In *Arabidopsis* roots, abscisic acid and gibberellin act antagonistically during root growth. Application of abscisic acid was demonstrated to stabilize the DELLA protein RGA and inhibit its gibberellin-induced degradation[94]. Furthermore, higher-order DELLA mutants are resistant to the effects of abscisic acid on growth inhibition[94]. DELLA proteins are not only a node for various hormone signal inputs but also mediate hormone signalling pathways in addition to gibberellin response. DELLA proteins were recently implicated as modulators of plant immune responses[95]. Genetic studies showed that DELLA proteins promote susceptibility to virulent biotrophs and resistance to necrotrophs. These observations were attributed, at least partially, to an alteration of the balance between salicylic acid and jasmonate signalling[95].

These are but a few of the common strategies used to coordinate hormone signals. The regulatory network connecting individual pathways is far more complex in that hormone interactions can result in different outcomes depending on the organ, developmental stage and environmental conditions. Continued efforts to describe the plant hormone regulatory network will be essential to understand and predict plant growth and development.

Concluding remarks

This is a very exciting time for plant hormone biologists as the rate of discovery has expanded exponentially in recent years. The number of recognized small-molecule hormones has doubled in the past 15 years. Candidate receptors for all of the 'classical' hormones and a few of the newer hormones have been identified. We now know the structure of several receptors, leading to exciting new models of hormone perception and new opportunities to synthesize novel growth regulators. Likewise, major advances have been achieved in our knowledge of downstream signalling components and some of their interactions. As we move forward, a major challenge will be to understand how hormone-signalling pathways are integrated during environmental control of plant growth. To take one example, auxin, gibberellin and the brassinosteroids are all known to promote elongation of the *Arabidopsis* hypocotyl. However, the relative contribution of each signalling pathway to growth regulation by light, temperature or the circadian clock is uncertain. Similarly, it is not known whether cell elongation in the hypocotyl requires the same set of genes, regardless of the growth signal. The answers to these and other questions will require a detailed characterization of growth responses, together with information on corresponding changes in the transcriptome and proteome, ideally at cellular resolution. Computational tools can be used to identify the gene modules associated with diverse growth responses. Ultimately, this information can be used to develop predictive models of plant growth and development that will be invaluable tools for modern agriculture.

1. Davies, P. J. in *Plant Hormones: Physiology, Biochemistry and Molecular Biology* (ed. Davies, P. J.) 1–12 (Kluwer Academic, 1995).
2. Grun, S., Lindermayr, C., Sell, S. & Durner, J. Nitric oxide and gene regulation in plants. *J. Exp. Bot.* **57**, 507–516 (2006).
3. Vert, G., Nemhauser, J. L., Geldner, N., Hong, F. & Chory, J. Molecular mechanisms of steroid hormone signaling in plants. *Annu. Rev. Cell Dev. Biol.* **21**, 177–201 (2005).
4. Browse, J. Jasmonate: an oxylipin signal with many roles in plants. *Vitam. Horm.* **72**, 431–456 (2005).
5. Loake, G. & Grant, M. Salicylic acid in plant defence—the players and protagonists. *Curr. Opin. Plant Biol.* **10**, 466–472 (2007).

6. Gomez-Roldan, V. *et al.* Strigolactone inhibition of shoot branching. *Nature* **455**, 189–194 (2008).
7. Umehara, M. *et al.* Inhibition of shoot branching by new terpenoid plant hormones. *Nature* **455**, 195–200 (2008).
References 6 and 7 were the first papers to identify strigolactones as plant hormones that play a major part in inhibiting axillary bud outgrowth.
8. Jun, J. H., Fiume, E. & Fletcher, J. C. The CLE family of plant polypeptide signaling molecules. *Cell. Mol. Life Sci.* **65**, 743–755 (2008).
9. Dharmasiri, N., Dharmasiri, S. & Estelle, M. The F-box protein TIR1 is an auxin receptor. *Nature* **435**, 441–445 (2005).
10. Dharmasiri, N. *et al.* Plant development is regulated by a family of auxin receptor F box proteins. *Dev. Cell* **9**, 109–119 (2005).
11. Kepinski, S. & Leyser, O. The *Arabidopsis* F-box protein TIR1 is an auxin receptor. *Nature* **435**, 446–451 (2005).
References 9 and 11 were the first papers to demonstrate that TIR1 was a receptor for auxin.
12. Ueguchi-Tanaka, M. *et al.* GIBBERELLIN INSENSITIVE DWARF1 encodes a soluble receptor for gibberellin. *Nature* **437**, 693–698 (2005).
13. Chini, A. *et al.* The JAZ family of repressors is the missing link in jasmonate signalling. *Nature* **448**, 666–671 (2007).
14. Thines, B. *et al.* JAZ repressor proteins are targets of the SCF(COI1) complex during jasmonate signalling. *Nature* **448**, 661–665 (2007).
References 13 and 14 were the first to demonstrate that the JAZ proteins were substrates of COI1 and that jasmonate promoted their interactions. This breakthrough strongly suggested that COI1 was a receptor for jasmonate and that the signalling pathway was mechanistically similar to auxin perception and signalling.
15. Melotto, M. *et al.* A critical role of two positively charged amino acids in the Jas motif of *Arabidopsis* JAZ proteins in mediating coronatine- and jasmonoyl isoleucine-dependent interactions with the COI1 F-box protein. *Plant J.* **55**, 979–988 (2008).
16. Pandey, S., Nelson, D. C. & Assmann, S. M. Two novel GPCR-type G proteins are abscisic acid receptors in *Arabidopsis. Cell* **136**, 136–148 (2009).
This paper identified two novel G-proteins that directly bind abscisic acid.
17. Chow, B. & McCourt, P. Plant hormone receptors: perception is everything. *Genes Dev.* **20**, 1998–2008 (2006).
18. Schwechheimer, C. & Willige, B. C. Shedding light on gibberellic acid signalling. *Curr. Opin. Plant Biol.* **12**, 57–62 (2009).
19. Rensing, S. A. *et al.* The *Physcomitrella* genome reveals evolutionary insights into the conquest of land by plants. *Science* **319**, 64–69 (2008).
20. Vandenbussche, F., Fierro, A. C., Wiedemann, G., Reski, R. & Van Der Straeten, D. Evolutionary conservation of plant gibberellin signalling pathway components. *BMC Plant Biol.* **7**, 65 (2007).
21. Zhao, Y. The role of local biosynthesis of auxin and cytokinin in plant development. *Curr. Opin. Plant Biol.* **11**, 16–22 (2008).
22. Mockaitis, K. & Estelle, M. Auxin receptors and plant development: a new signaling paradigm. *Annu. Rev. Cell Dev. Biol.* **24**, 55–80 (2008).
23. Wasternack, C. Jasmonates: an update on biosynthesis, signal transduction and action in plant stress response, growth and development. *Ann. Bot. (Lond.)* **100**, 681–697 (2007).
24. Yamaguchi, S. Gibberellin metabolism and its regulation. *Annu. Rev. Plant Biol.* **59**, 225–251 (2008).
25. Hirano, K., Ueguchi-Tanaka, M. & Matsuoka, M. GID1-mediated gibberellin signaling in plants. *Trends Plant Sci.* **13**, 192–199 (2008).
26. Hirose, N. *et al.* Regulation of cytokinin biosynthesis, compartmentalization and translocation. *J. Exp. Bot.* **59**, 75–83 (2008).
27. Symons, G. M., Ross, J. J., Jager, C. E. & Reid, J. B. Brassinosteroid transport. *J. Exp. Bot.* **59**, 17–24 (2008).
28. Hirayama, T. & Shinozaki, K. Perception and transduction of abscisic acid signals: keys to the function of the versatile plant hormone ABA. *Trends Plant Sci.* **12**, 343–351 (2007).
29. Dharmasiri, S. & Estelle, M. The role of regulated protein degradation in auxin response. *Plant Mol. Biol.* **49**, 401–409 (2002).
30. Moon, J., Parry, G. & Estelle, M. The ubiquitin-proteasome pathway and plant development. *Plant Cell* **16**, 3181–3195 (2004).
31. Ruegger, M. *et al.* The TIR1 protein of *Arabidopsis* functions in auxin response and is related to human SKP2 and yeast grr1p. *Genes Dev.* **12**, 198–207 (1998).
32. Gray, W. M., Kepinski, S., Rouse, D., Leyser, O. & Estelle, M. Auxin regulates SCFTIR1-dependent degradation of AUX/IAA proteins. *Nature* **414**, 271–276 (2001).
33. Guilfoyle, T. J. & Hagen, G. Auxin response factors. *Curr. Opin. Plant Biol.* **10**, 453–460 (2007).
34. Reed, J. W. Roles and activities of Aux/IAA proteins in *Arabidopsis. Trends Plant Sci.* **6**, 420–425 (2001).
35. Dharmasiri, N., Dharmasiri, S., Jones, A. M. & Estelle, M. Auxin action in a cell-free system. *Curr. Biol.* **13**, 1418–1422 (2003).
36. Kepinski, S. & Leyser, O. Auxin-induced SCFTIR1-Aux/IAA interaction involves stable modification of the SCFTIR1 complex. *Proc. Natl Acad. Sci. USA* **101**, 12381–12386 (2004).
37. Tan, X. *et al.* Mechanism of auxin perception by the TIR1 ubiquitin ligase. *Nature* **446**, 640–645 (2007).
This paper described the crystal structure of the TIR1 auxin receptor in complex with auxin and its substrate. The structure has substantially improved our understanding of auxin perception.
38. Tan, X. & Zheng, N. Hormone signaling through protein destruction: a lesson from plants. *Am. J. Physiol. Endocrinol. Metab.* **296**, E223–E227 (2009).
39. Hayashi, K. *et al.* Small-molecule agonists and antagonists of F-box protein-substrate interactions in auxin perception and signaling. *Proc. Natl Acad. Sci. USA* **105**, 5632–5637 (2008).
40. Tiwari, S. B., Hagen, G. & Guilfoyle, T. J. Aux/IAA proteins contain a potent transcriptional repression domain. *Plant Cell* **16**, 533–543 (2004).
41. Szemenyei, H., Hannon, M. & Long, J. A. TOPLESS mediates auxin-dependent transcriptional repression during *Arabidopsis* embryogenesis. *Science* **319**, 1384–1386 (2008).
This paper demonstrated that TOPLESS directly binds domain I of Aux/IAA proteins, leading to transcriptional repression of auxin-responsive genes. This mechanism probably accounts for the repression of ARF function by Aux/IAA proteins.
42. Feys, B., Benedetti, C. E., Penfold, C. N. & Turner, J. G. *Arabidopsis* mutants selected for resistance to the phytotoxin coronatine are male sterile, insensitive to methyl jasmonate, and resistant to a bacterial pathogen. *Plant Cell* **6**, 751–759 (1994).
43. Xie, D. X., Feys, B. F., James, S., Nieto-Rostro, M. & Turner, J. G. COI1: an *Arabidopsis* gene required for jasmonate-regulated defense and fertility. *Science* **280**, 1091–1094 (1998).
44. Yan, Y. *et al.* A downstream mediator in the growth repression limb of the jasmonate pathway. *Plant Cell* **19**, 2470–2483 (2007).
45. Katsir, L., Schilmiller, A. L., Staswick, P. E., He, S. Y. & Howe, G. A. COI1 is a critical component of a receptor for jasmonate and the bacterial virulence factor coronatine. *Proc. Natl Acad. Sci. USA* **105**, 7100–7105 (2008).
46. de Lucas, M. *et al.* A molecular framework for light and gibberellin control of cell elongation. *Nature* **451**, 480–484 (2008).
47. Feng, S. *et al.* Coordinated regulation of *Arabidopsis thaliana* development by light and gibberellins. *Nature* **451**, 475–479 (2008).
References 46 and 47 revealed that the basic helix–loop–helix transcription factors PIF3 and PIF4 directly interact with DELLA proteins. Upon gibberellin accumulation, the DELLA proteins are destabilized, freeing PIF3 and PIF4 to regulate their target genes. This is a new and compelling model for DELLA-mediated growth regulation.
48. McGinnis, K. M. *et al.* The *Arabidopsis* SLEEPY1 gene encodes a putative F-box subunit of an SCF E3 ubiquitin ligase. *Plant Cell* **15**, 1120–1130 (2003).
49. Sasaki, A. *et al.* Accumulation of phosphorylated repressor for gibberellin signaling in an F-box mutant. *Science* **299**, 1896–1898 (2003).
50. Nakajima, M. *et al.* Identification and characterization of *Arabidopsis* gibberellin receptors. *Plant J.* **46**, 880–889 (2006).
51. Griffiths, J. *et al.* Genetic characterization and functional analysis of the GID1 gibberellin receptors in *Arabidopsis. Plant Cell* **18**, 3399–3414 (2006).
52. Willige, B. C. *et al.* The DELLA domain of GA INSENSITIVE mediates the interaction with the GA INSENSITIVE DWARF1A gibberellin receptor of *Arabidopsis. Plant Cell* **19**, 1209–1220 (2007).
53. Murase, K., Hirano, Y., Sun, T. P. & Hakoshima, T. Gibberellin-induced DELLA recognition by the gibberellin receptor GID1. *Nature* **456**, 459–463 (2008).
54. Shimada, A. *et al.* Structural basis for gibberellin recognition by its receptor GID1. *Nature* **456**, 520–523 (2008).
References 53 and 54 are the first to describe the structure of the GID1 gibberellin receptor.
55. Shen, Y. Y. *et al.* The Mg-chelatase H subunit is an abscisic acid receptor. *Nature* **443**, 823–826 (2006).
56. Mochizuki, N., Brusslan, J. A., Larkin, R., Nagatani, A. & Chory, J. *Arabidopsis* genomes uncoupled 5 (GUN5) mutant reveals the involvement of Mg-chelatase H subunit in plastid-to-nucleus signal transduction. *Proc. Natl Acad. Sci. USA* **98**, 2053–2058 (2001).
57. Muller, A. H. & Hansson, M. The barley magnesium chelatase 150-kDa subunit is not an abscisic-acid receptor. *Plant Physiol.* **150**, 157–166 (2009).
58. Liu, X. *et al.* A G protein-coupled receptor is a plasma membrane receptor for the plant hormone abscisic acid. *Science* **315**, 1712–1716 (2007).
59. Johnston, C. A. *et al.* Comment on "A G protein coupled receptor is a plasma membrane receptor for the plant hormone abscisic acid". *Science* **318**, 914; author reply 914 (2007).
60. Guo, J., Zeng, Q., Emami, M., Ellis, B. E. & Chen, J. G. The GCR2 gene family is not required for ABA control of seed germination and early seedling development in *Arabidopsis. PLoS One* **3**, e2982 (2008).
61. Gao, Y. *et al.* Genetic characterization reveals no role for the reported ABA receptor, GCR2, in ABA control of seed germination and early seedling development in *Arabidopsis. Plant J.* **52**, 1001–1013 (2007).
62. Risk, J. M., Day, C. L. & Macknight, R. C. Re-evaluation of abscisic acid (ABA) binding assays shows that GCR2 does not bind ABA. *Plant Physiol.* **150**, 6–11 (2009).
63. Park, S. Y. *et al.* Abscisic acid inhibits type 2C protein phosphatases via the PYR/PYL family of START proteins. *Science* Epub ahead of print, doi:10.1126/science.1173041 (30 April 2009).
64. Ma, Y. *et al.* Regulators of PP2C phosphatase activity function as abscisic acid sensors. *Science* Epub ahead of print, doi:10.1126/science.1172408 (8 May 2009).

References 63 and 64 describe the identification of soluble abscisic-acid receptors. A novel model for abscisic-acid action is proposed, in which abscisic acid acts to inhibit PP2C proteins such as ABI1 and ABI2, by promoting an interaction between the phosphatase and PYR1/RCAR1 and related proteins.
65. McCourt, P. & Creelman, R. The ABA receptors—we report, you decide. *Curr. Opin. Plant Biol.* **11,** 474–478 (2008).
66. Zhang, X., Garreton, V. & Chua, N. H. The AIP2 E3 ligase acts as a novel negative regulator of ABA signaling by promoting ABI3 degradation. *Genes Dev.* **19,** 1532–1543 (2005).
67. Stone, S. L., Williams, L. A., Farmer, L. M., Vierstra, R. D. & Callis, J. KEEP ON GOING, a RING E3 ligase essential for *Arabidopsis* growth and development, is involved in abscisic acid signaling. *Plant Cell* **18,** 3415–3428 (2006).
68. Solano, R., Stepanova, A., Chao, Q. & Ecker, J. R. Nuclear events in ethylene signaling: a transcriptional cascade mediated by ETHYLENE-INSENSITIVE3 and ETHYLENE-RESPONSE-FACTOR1. *Genes Dev.* **12,** 3703–3714 (1998).
69. Guo, H. & Ecker, J. R. Plant responses to ethylene gas are mediated by SCF(EBF1/EBF2)-dependent proteolysis of EIN3 transcription factor. *Cell* **115,** 667–677 (2003).
70. Potuschak, T. *et al.* EIN3-dependent regulation of plant ethylene hormone signaling by two *Arabidopsis* F box proteins: EBF1 and EBF2. *Cell* **115,** 679–689 (2003).
References 67 and 68 showed that the F-box proteins EBF1 and EBF2 were important regulators of EIN3 accumulation. The data demonstrate the pivital role of the ubiquitin-mediated degradation during ethylene signalling.
71. Binder, B. M. *et al.* The *Arabidopsis* EIN3 binding F-box proteins EBF1 and EBF2 have distinct but overlapping roles in ethylene signaling. *Plant Cell* **19,** 509–523 (2007).
72. Qiao, H., Chang, K. N., Yazaki, J. & Ecker, J. R. Interplay between ethylene, ETP1/ETP2 F-box proteins, and degradation of EIN2 triggers ethylene responses in *Arabidopsis. Genes Dev.* **23,** 512–521 (2009).
73. Ongaro, V. & Leyser, O. Hormonal control of shoot branching. *J. Exp. Bot.* **59,** 67–74 (2008).
74. Stirnberg, P., van De Sande, K. & Leyser, H. M. MAX1 and MAX2 control shoot lateral branching in *Arabidopsis. Development* **129,** 1131–1141 (2002).
75. Johnson, X. *et al.* Branching genes are conserved across species. Genes controlling a novel signal in pea are coregulated by other long-distance signals. *Plant Physiol.* **142,** 1014–1026 (2006).
76. Wilson, A. K., Pickett, F. B., Turner, J. C. & Estelle, M. A dominant mutation in *Arabidopsis* confers resistance to auxin, ethylene and abscisic acid. *Mol. Gen. Genet.* **222,** 377–383 (1990).
77. Roman, G., Lubarsky, B., Kieber, J. J., Rothenberg, M. & Ecker, J. R. Genetic analysis of ethylene signal transduction in *Arabidopsis thaliana*: five novel mutant loci integrated into a stress response pathway. *Genetics* **139,** 1393–1409 (1995).
78. Stepanova, A. N., Hoyt, J. M., Hamilton, A. A. & Alonso, J. M. A link between ethylene and auxin uncovered by the characterization of two root-specific ethylene-insensitive mutants in *Arabidopsis. Plant Cell* **17,** 2230–2242 (2005).
79. Stepanova, A. N. *et al.* TAA1-mediated auxin biosynthesis is essential for hormone crosstalk and plant development. *Cell* **133,** 177–191 (2008).
80. Tao, Y. *et al.* Rapid synthesis of auxin via a new tryptophan-dependent pathway is required for shade avoidance in plants. *Cell* **133,** 164–176 (2008).
81. Tsuchisaka, A. & Theologis, A. Unique and overlapping expression patterns among the *Arabidopsis* 1-amino-cyclopropane-1-carboxylate synthase gene family members. *Plant Physiol.* **136,** 2982–3000 (2004).
82. Tsuchisaka, A. & Theologis, A. Heterodimeric interactions among the 1-amino-cyclopropane-1-carboxylate synthase polypeptides encoded by the *Arabidopsis* gene family. *Proc. Natl Acad. Sci. USA* **101,** 2275–2280 (2004).
83. Nagpal, P. *et al.* Auxin response factors ARF6 and ARF8 promote jasmonic acid production and flower maturation. *Development* **132,** 4107–4118 (2005).
84. Feraru, E. & Friml, J. PIN polar targeting. *Plant Physiol.* **147,** 1553–1559 (2008).
85. Laplaze, L. *et al.* Cytokinins act directly on lateral root founder cells to inhibit root initiation. *Plant Cell* **19,** 3889–3900 (2007).
86. Blilou, I. *et al.* The PIN auxin efflux facilitator network controls growth and patterning in *Arabidopsis* roots. *Nature* **433,** 39–44 (2005).
87. Nemhauser, J. L., Mockler, T. C. & Chory, J. Interdependency of brassinosteroid and auxin signaling in *Arabidopsis. PLoS Biol.* **2,** E258 (2004).
88. Goda, H. *et al.* Comprehensive comparison of auxin-regulated and brassinosteroid-regulated genes in *Arabidopsis. Plant Physiol.* **134,** 1555–1573 (2004).
89. Mouchel, C. F., Osmont, K. S. & Hardtke, C. S. BRX mediates feedback between brassinosteroid levels and auxin signalling in root growth. *Nature* **443,** 458–461 (2006).
90. Vert, G., Walcher, C. L., Chory, J. & Nemhauser, J. L. Integration of auxin and brassinosteroid pathways by Auxin Response Factor 2. *Proc. Natl Acad. Sci. USA* **105,** 9829–9834 (2008).
This paper elucidated an interesting molecular crosstalk strategy in which the brassinosteroid-regulated BIN2 kinase directly regulates ARF2, thereby modulating auxin signalling.
91. Weiss, D. & Ori, N. Mechanisms of cross talk between gibberellin and other hormones. *Plant Physiol.* **144,** 1240–1246 (2007).
92. Fu, X. & Harberd, N. P. Auxin promotes *Arabidopsis* root growth by modulating gibberellin response. *Nature* **421,** 740–743 (2003).
93. Achard, P., Vriezen, W. H., Van Der Straeten, D. & Harberd, N. P. Ethylene regulates *Arabidopsis* development via the modulation of DELLA protein growth repressor function. *Plant Cell* **15,** 2816–2825 (2003).
94. Achard, P. *et al.* Integration of plant responses to environmentally activated phytohormonal signals. *Science* **311,** 91–94 (2006).
95. Navarro, L. *et al.* DELLAs control plant immune responses by modulating the balance of jasmonic acid and salicylic acid signaling. *Curr. Biol.* **18,** 650–655 (2008).

Acknowledgements Work in M.E.'s laboratory was supported by grants from the NIH (GM43644), the NSF (IOS 0744800), and the DOE (DOE DE-FG02-02ER15312) to M.E.

Author Contributions A.S. and M.E. generated an outline together. A.S. prepared the first draft of the article and A.S. and M.E. worked together on all subsequent drafts.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence should be addressed to M.E. (mestelle@ucsd.edu).

ARTICLES

A soma-to-germline transformation in long-lived *Caenorhabditis elegans* mutants

Sean P. Curran[1,2], Xiaoyun Wu[1,2], Christian G. Riedel[1,2] & Gary Ruvkun[1,2]

Unlike the soma, which ages during the lifespan of multicellular organisms, the germ line traces an essentially immortal lineage. Genomic instability in somatic cells increases with age, and this decline in somatic maintenance might be regulated to facilitate resource reallocation towards reproduction at the expense of cellular senescence. Here we show that *Caenorhabditis elegans* mutants with increased longevity exhibit a soma-to-germline transformation of gene expression programs normally limited to the germ line. Decreased insulin-like signalling causes the somatic misexpression of the germline-limited *pie-1* and *pgl* family of genes in intestinal and ectodermal tissues. The forkhead boxO1A (FOXO) transcription factor DAF-16, the major transcriptional effector of insulin-like signalling, regulates *pie-1* expression by directly binding to the *pie-1* promoter. The somatic tissues of insulin-like mutants are more germline-like and protected from genotoxic stress. Gene inactivation of components of the cytosolic chaperonin complex that induce increased longevity also causes somatic misexpression of PGL-1. These results indicate that the acquisition of germline characteristics by the somatic cells of *C. elegans* mutants with increased longevity contributes to their increased health and survival.

In genetic and functional genomic screens for gene inactivations that increase the lifespan of *C. elegans*, decreases in the insulin-like signalling pathway confer the largest increase in longevity. Inactivation of essential genes that mediate core functions, such as RNA transcription and splicing, translation initiation and chromatin-remodelling, constitute the second tier of potent longevity enhancing pathways[1]. In addition to regulating longevity, mutations in the insulin receptor homologue *daf-2*, as well as mutations in downstream signalling components such as the phosphatidylinositol-3-kinase (PI3K) *age-1*, and the serine/threonine kinases *akt-1* and *pdk-1*, exhibit enhanced immune responses to bacterial pathogens[2] and enhanced RNA interference (RNAi) responses in somatic cells[3]. Because RNAi constitutes an immune response in *C. elegans*[4], it is reasonable that the stress-resistance pathway triggered by low insulin-like signalling would intensify RNAi. The molecular basis of how decreased *daf-2* signalling enhances RNAi is unknown, but it is mediated by the major transcriptional output of insulin-like signalling, the FOXO transcription factor DAF-16, because mutations in *daf-16* suppress the enhanced RNAi of insulin-signalling-deficient mutants[3]. Another important pathway that negatively regulates *C. elegans* RNAi is the class B synthetic multivulva (synMuvB) pathway of chromatin factors that also regulate the repression of germline fates in somatic cells[5]. Given that the germline genome is much more protected than the soma from viral invasion and mutational events by both small-RNA-mediated and other genome surveillance pathways[6–9], we explored whether the enhanced RNAi phenotypes of insulin signalling mutants and other long-lived mutations engage germline versus somatic cell specification programs.

Mutants in the synMuvB class of genes, which includes the *lin-35* retinoblastoma homologue, show an enhanced response to double-stranded (ds)RNA and somatic misexpression of the normally germline-limited P-granule component PGL-1 (ref. 5). This expression of germline genes in somatic tissues has been linked to the enhanced RNAi phenotype, because germline cells are remarkably proficient at RNAi. *pgl-1* encodes an RNA-binding protein that is localized to P granules and is essential for gene silencing by cosuppression in the germ line[10], supporting the model that somatic misexpression of *pgl-1* contributes to enhanced somatic cell RNAi. Thus, increased longevity mutants and germ cells both have exceptional longevity, are resistant to cellular stress, and are enhanced for RNAi. If a soma-to-germline transformation phenotype were present in the insulin-like signalling mutants, then the germ line quality of somatic cells could explain the enhanced RNAi phenotype and potentially contribute to the remarkable lifespan extension (Supplementary Fig. 1).

Germ cell character in somatic tissues

We tested whether the insulin-like-signalling mutants induce a soma-to-germline transformation of cell fate. A protein fusion (GFP::PGL-1), containing green fluorescent protein (GFP), expressed under the control of the germline-specific *pie-1* promoter was monitored in *daf-2*(*e1370*) (the insulin-like receptor) or *age-1*(*mg305*) (the downstream PI3K) mutant strains. *pie-1* and *pgl-1* are exclusively expressed in the germ line of wild-type animals (Fig. 1a–c and Supplementary Fig. 2)[11,12]. Decreased insulin-like signalling caused strong misexpression of GFP::PGL-1 in hypodermal and intestinal somatic tissues of dauers (Fig. 1d, e) and late larval stage animals (Fig. 1f, g and Supplementary Fig. 2).

To monitor the *in vivo* regulation of endogenous *pie-1* transcript levels, we performed quantitative PCR with reverse transcription (qRT–PCR) from *daf-2*(*e1370*) and *daf-2*(*e1368*) dauers. This stage was chosen because GFP::PGL-1 expression from the *pie-1* promoter was the strongest in dauers (Fig. 1d), and germline transcripts would be attenuated because the dauer germ line is less developed (Table 1, Supplementary Table 1 and Supplementary Figs 5 and 6)[13]. Comparison of both *daf-2*(*e1368*) and *daf-2*(*e1370*) dauers to age-matched wild-type larval stage 2/3 larvae showed a 2.1 ± 0.1-fold and 4.5 ± 1.0-fold increase in *pie-1* transcript levels, respectively (Table 1 and Supplementary Table 1). DAF-16 is the most downstream transcriptional effector of the insulin-like signalling pathway[14,15], and its nuclear localization is activated in low insulin-like

[1]Department of Molecular Biology, Massachusetts General Hospital, Boston, Massachusetts 02114, USA. [2]Department of Genetics, Harvard Medical School, Boston, Massachusetts 02114, USA.

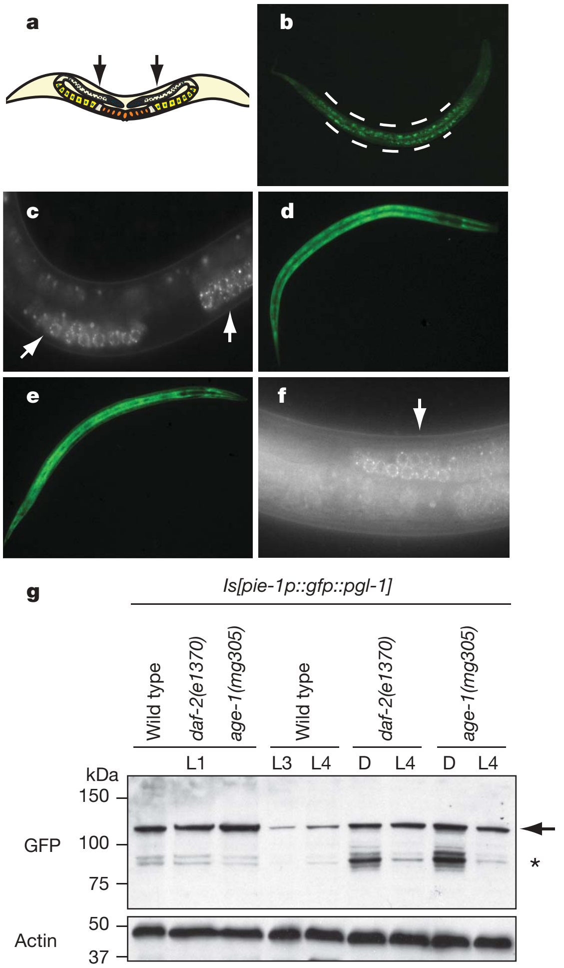


Figure 1 | Mutations in the insulin/IGF-like signalling pathway cause soma-to-germline transformation. **a**, Schematic of the *C. elegans* germ line. **b**, **c**, Wild-type larval stage 2/3 larvae express *pie-1p::gfp::pgl-1* only in the germ line in perinuclear structures (arrows in **c**, **f**). The region between the dashed lines in **b** denotes background autofluorescence. **d**–**f**, *daf-2*(*e1370*) (**d**) and *age-1*(*mg305*) (**e**) mutations cause somatic expression of *pie-1p::gfp::pgl-1* in the intestine and hypodermis in dauers and *daf-2*(*e1370*) larval stage 3/4 larvae (**f**). Original magnifications, ×10 (**b**, **d**, **e**) and ×40 (**c**, **f**). **g**, Western analysis of SDS–PAGE-resolved whole worm lysates of wild-type, *daf-2*(*e1370*), and *age-1*(*mg305*) animals containing the *pie-1p::gfp::pgl-1* reporter. Arrow denotes GFP::PGL-1, asterisk marks an unknown crossreacting protein. D, dauer; L1, larval stage 1; L3, larval stage 3; L4, larval stage 4.

signalling conditions. Loss of DAF-16 abrogates the longevity and enhanced RNAi phenotypes caused by defects in insulin-like signalling[3,14,16]. In agreement, the *daf-16*(*mgDf47*); *daf-2*(*e1370*) double mutant had reduced *pie-1* transcript levels (Table 1 and Supplementary Table 1). We looked for *pie-1* misexpression in other longevity promoting manipulations that function independently of insulin-like signalling, such as mitochondrial *clk-1* RNAi[17], and dietary restriction via an *eat-2*(*ad465*)[18] mutation and food deprivation[19] (Supplementary Fig. 7). None of these measures induced the misexpression of germline genes observed in the insulin-like signalling mutants.

There are potential DAF-16-binding sites in the *pie-1* promoter region, predicting that the misexpression of the germline-specific transgene and endogenous *pie-1* transcript would occur at the level of DAF-16 binding to the *pie-1* promoter element. We examined the ability of purified recombinant DAF-16 (Supplementary Fig. 3) to bind to regions of the *pie-1* promoter that contain predicted DAF-16-binding/associated elements by electrophoretic mobility gel shift assays (EMSA) (Fig. 2a, b)[20–22]. DAF-16 was capable of binding to most fragments of the *pie-1* promoter tested (Fig. 2b, lane 2 and Supplementary Fig. 4). This binding was specific to the purified DAF-16 protein, as shown by supershift analysis with a monoclonal antibody that recognizes a Flag epitope tag engineered into the carboxy terminus of the recombinant protein (Fig. 2b, lane 3), and by competition of binding with a cold dsDNA oligonucleotide containing several DAF-16-associated elements (Fig. 2b, lane 4).

Under conditions of low insulin-like signalling, DAF-16 has been shown to occupy the promoters of stress resistance and detoxifying enzymes such as *sod-3* (ref. 22) and *mtl-1* (ref. 23). We tested for *in vivo* binding of DAF-16 to the *pie-1* promoter using chromatin immunoprecipitation (ChIP) with a DAF-16::GFP fusion gene that is only expressed in somatic tissues. To induce DAF-16 nuclear localization,

Table 1 | Insulin-like signalling regulates the expression of germline genes in somatic cells

	pie-1	*pgl-1*	*pgl-2*	*pgl-3*
*daf-2(e1368)**	2.1 ± 0.1§	1.0 ± 0.1	4.6 ± 0.9‡	1.3 ± 0.2‡
*daf-2(e1370)**	4.5 ± 1.0‡	1.4 ± 0.2	3.3 ± 0.2§	1.9 ± 0.3§
*daf-16(mgDf47); daf-2(e1370)**	0.8 ± 0.1‡	0.4 ± 0§	0.9 ± 0.1	0.6 ± 0.1§
*glp-4(bn2)**	0.007 ± 0§	0.007 ± 0§	0.005 ± 0§	0.005 ± 0§
glp-4(bn2); daf-2(e1370)†	3.7 ± 1.0§	3.9 ± 1.5§	2.8 ± 0.1§	2.3 ± 0.8§

Values indicate mean and s.d., *n* = ≥3 biological replicates. See Supplementary Table 1 for raw data sets.
* qRT–PCR analysis is represented as the fold change relative to wild-type samples normalized to *rpl-32* and *snb-1* expression.
† Samples are compared to *glp-4*(*bn2*) mRNA levels normalized to *rpl-32* and *snb-1* expression.
‡ $P < 0.05$, Student's *t*-test.
§ $P < 0.01$, Student's *t*-test.

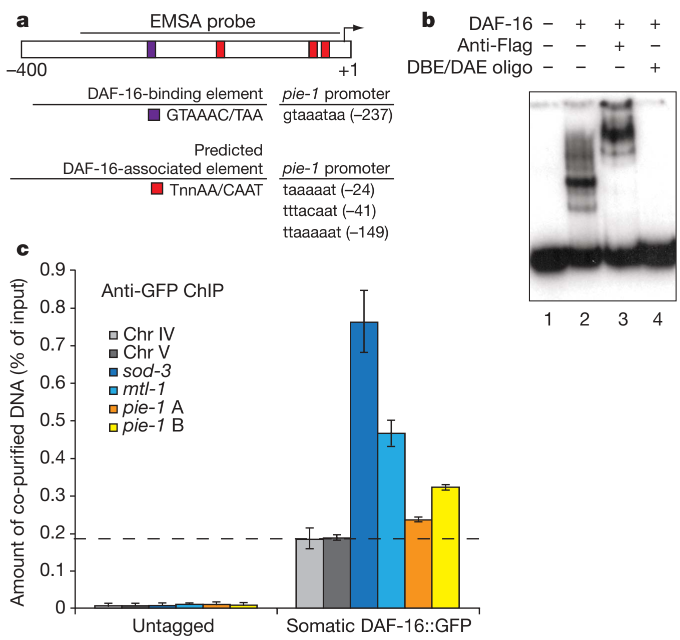


Figure 2 | DAF-16 regulates the expression of *pie-1*. **a**, Schematic of the *pie-1* promoter. **b**, A ^{32}P-labelled region of the *pie-1* promoter (EMSA probe, lane 1) can recruit purified recombinant DAF-16–Flag (lane 2). The bound DAF-16 protein can be supershifted with anti-Flag antibodies (lane 3). Binding is competed away with a cold dsDNA oligonucleotide containing DAF-16-binding/associated elements (DBE/DAE, lane 4). **c**, DAF-16 is enriched on the promoter of *pie-1* in *daf-2*-RNAi-treated worms. ChIP was performed using anti-GFP antibodies in wild type (untagged) and a somatic exclusive *daf-16::gfp* strain (somatic DAF-16::GFP). The binding of two non-DAF-16-associated control regions (Chr IV, Chr V), two DAF-16-associated promoter regions (*sod-3*, *mtl-1*), and two regions in the *pie-1* promoter (*pie-1* A, *pie-1* B) was measured from the immunoprecipitation. Error bars represent the s.e.m. of duplicated reactions. Similar results were obtained from at least two independent experiments with two different tagged DAF-16 strains (see Supplementary Table 3).

we inactivated *daf-2* by RNAi. Notably, we found significant enrichment of DAF-16 at the somatic *pie-1* promoter (Fig. 2c).

The finding that *pie-1* messenger RNA is increased during dauer diapause correlates with the observations that the transcriptional activity of dauers is reduced by ~85% compared to other larval stages, and that dauers are exceptionally long-lived[24]. PIE-1 functions as a global repressor of RNA polymerase II transcription in the germ line during early embryonic development[25], and could have a similar role in maintaining a transcriptionally repressed state in the soma of a dauer animal.

Ectopic somatic expression of *pie-1* causes misexpression of germline genes in somatic cells, a phenotype very similar to the misexpression of germline genes induced in the synMuvB class mutants[5,26]. Analysis by qRT–PCR of the *pgl* family of germline-limited genes showed that they too were expressed at higher levels in *daf-2* mutants in a *daf-16*-dependent manner (Table 1 and Supplementary Table 1). Because the endogenous mRNA contribution from a larval stage 3 germ line is two to threefold higher than a dauer germ line[13], this analysis could systematically underestimate the extent of soma-to-germline transformation of cell fate in the *daf-2* mutants. To eliminate germline contribution to these whole animal mRNA assays, we genetically ablated the germ line of a *daf-2* mutant using a *glp-4* germline proliferation defective mutation[27], and analysed the expression of these germline-specific genes by qRT–PCR. Comparison of *pie-1*, *pgl-1*, *pgl-2* and *pgl-3* expression levels between wild type and *glp-4* shows that these genes are germline-specific in wild-type animals and therefore expressed at a far lower level in the *glp-4* mutant (Table 1). The *glp-4* mutant has little or no lifespan change relative to wild type[28]. The somatic expression of *pie-1*, *pgl-1*, *pgl-2* and *pgl-3* transcripts was much more notable comparing *glp-4; daf-2* animals to the *glp-4* control (Table 1).

Somatic protection of insulin-like signalling mutants

The somatic misexpression of germline genes suggests that the somatic cells of an insulin-like-signalling mutant are more germline-like, and this transformation may have functional consequences. The *C. elegans* germ line is exceptionally proficient at RNAi and perhaps for this reason, the somatic tissues of insulin-like signalling mutants are enhanced for RNAi[3]. We proposed that germline-transformed somatic cells, like germ cells, may engage further protective pathways that prevent or slow genomic destabilization, which could increase the ability to respond to stress and extend lifespan[29,30].

We tested the ability of *daf-2*(*e1368*) mutants to protect somatic tissues from genomic instability by RNAi-depleting genes previously shown to protect the genome[31,32]. Gene inactivation of *rpa-2* (WormBase sequence name *M04F3.1*), *srxa-6* (*W07A8.5*), *dnJ-25* (*W07A8.3*) or F49E12.6 caused an increased spontaneous mutation rate and reduced viability in the wild type, presumably owing to increased somatic mutation (Supplementary Fig. 8). *daf-2*(*e1368*) animals fed the same RNAi clones showed much lower levels of developmental arrest than wild-type animals, suggesting that the *daf-2* mutants were better protected from this form of genotoxic stress than the wild-type animals (Supplementary Fig. 8). These same gene inactivations in the wild type induce DNA damage and DNA strand slippage of a tandem repeat, as measured by the somatic expression of a normally out-of-frame lacZ-GFP transgene[32]. Inactivation of *srxa-6* and *dnJ-25* by RNAi causes strong intestinal expression of the out-of-frame lacZ-fusion gene in wild-type animals, whereas far fewer *daf-2*(*e1368*) mutant animals re-animated this out-of-frame fusion gene, suggesting a reduction in genomic instability in the somatic cells of the insulin-like signalling mutants (Supplementary Fig. 9). The fact that *daf-2* mutants are enhanced for RNAi suggests that the resistance to somatic DNA damage is not a consequence of an attenuated response to the RNAi treatment, but is instead an increased ability of the somatic cells to respond to deficits in pathways that regulate genomic surveillance.

Soma-to-germline transformation promotes cell survival

If the misexpression of germline genes in somatic tissues contributes to the longevity phenotype of insulin-like signalling mutants, then the inactivation of these genes by RNAi should suppress the longevity phenotype. To test this we inactivated several germline genes misexpressed in the somatic cells of *daf-2* mutants. To rule out pleiotropic effects, we also performed a similar analysis in wild-type animals. Depletion of these misexpressed germline genes in the *daf-2* mutant reduced the health of these normally long-lived animals. The same gene inactivations in wild-type animals caused a slight increase in lifespan (Fig. 3 and Supplementary Table 2). This phenotype is probably caused by the reduction in the proliferating germ-cell population within the gonad, which has been previously shown to increase lifespan in wild-type animals[33]. Although none of the gene inactivations could suppress *daf-2* lifespan as well as inactivation of *daf-16*, it is interesting that these gene inactivations changed sign for longevity control in this genetic background compared to wild type. The somatic expression of germline features in the synMuvB mutants is suppressed by loss-of-function mutations in the chromatin remodelling ATPase, *isw-1* and the SET-domain-containing protein *mes-4* (ref. 34). Similarly, RNAi inactivation of *isw-1* and *mes-4* could partially suppress the enhanced longevity of *daf-2* mutants.

Mechanism of soma-to-germline transformation

An RNAi screen for essential gene inactivations that increase lifespan identified a large number of longevity genes that regulate gene expression and possibly RNAi. We tested whether any of the other genes identified in our post-developmental RNAi screen to extend lifespan also showed expression of germline markers in somatic tissues by tracking the expression of endogenous PGL-1 (ref. 1). PGL-1 is normally restricted to the germ line where it forms perinuclear punctate structures around mitotic and meiotic germ cell nuclei (Fig. 4). Gene

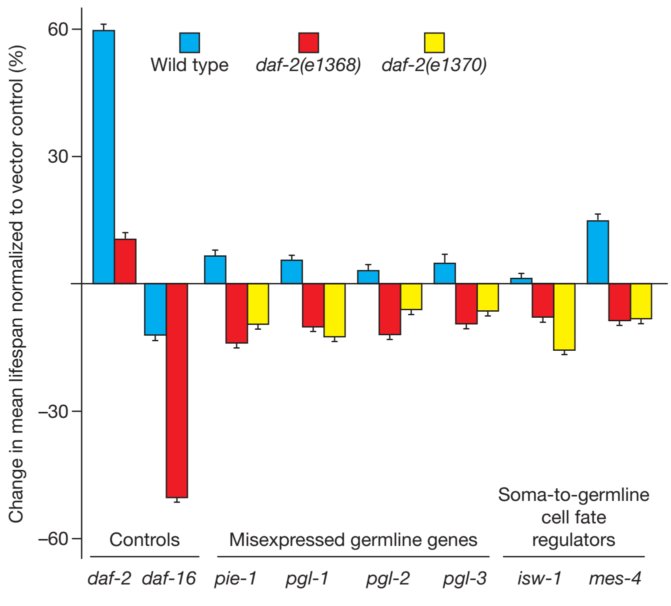


Figure 3 | The somatic misexpression of germline-specific genes in insulin-like signalling mutants contributes to their increased longevity. Wild-type animals (blue) or two hypomorphic insulin-like receptor mutants *daf-2*(*e1368*) (red) and *daf-2*(*e1370*) (yellow) were fed RNAi clones targeting germline-specific genes that are misexpressed in somatic cells and two known soma/germline-specificity regulators only in adulthood. RNAi depletion of germline-restricted genes in wild-type animals increases lifespan compared to vector control, whereas inactivation in *daf-2* mutants shortens the lifespan of these normally long-lived animals. RNAi targeting *daf-2* (increased lifespan) or *daf-16* (decreased lifespan) were used as controls, and lifespan was normalized to animals fed vector control RNAi. Error bars represent s.e.m. (see Supplementary Table 2 for numbers and detailed statistical analysis).

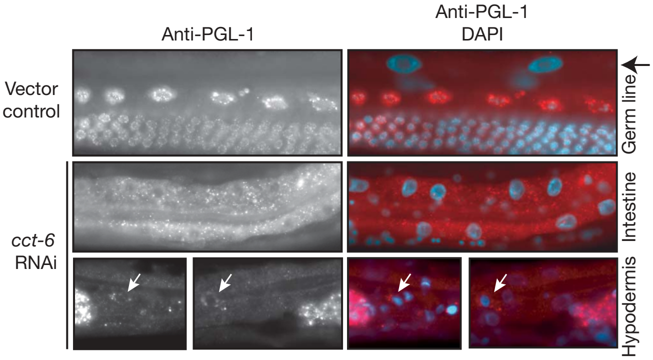


Figure 4 | The cytosolic chaperonin complex regulates the expression of PGL-1 in somatic cells. PGL-1 immunostaining is restricted to perinuclear structures of meiotic and mitotic germ cells in animals fed vector control (top panel). The black arrow points to four intestinal nuclei lacking PGL-1 expression. RNAi targeting *cct-6* causes somatic PGL-1 expression in intestinal (middle panel) and hypodermal cells (bottom panel). White arrows indicate perinuclear PGL-1 localization. DAPI, 4,6-diamidino-2-phenylindole. Original magnification, ×40.

inactivation by RNAi of two components of the cytosolic chaperonin complex *cct-4* and *cct-6* that increase longevity, caused somatic expression of PGL-1 as detected by antibody staining (Fig. 4 and Supplementary Fig. 10). The somatic PGL-1 protein localized to punctate perinuclear rings in hypodermal cells and also produced intestinal cytoplasmic granules. It was notable that these are the same tissues in which we observed misexpression of *pie-1* in the insulin-like signalling mutants.

The cytosolic chaperonin complex negative regulates the activity of intestinal *skn-1* (ref. 35). Loss of cytosolic chaperonin complex function causes SKN-1-dependent transcriptional activation of detoxifying enzymes and stress response gene targets such as *gst-4* (refs 35, 36). Furthermore, intestinal *skn-1* mediates the longevity phenotype of insulin-like signalling mutants parallel to *daf-16* (ref. 37). In support of the intestine acting as an important tissue for this regulation, the strongest misexpression of PGL-1 was detected in the intestine of *cct-4* and *cct-6* RNAi-treated animals. Together these data support a model in which *C. elegans* longevity-promoting mutations, such as decreased insulin-like signalling or reduction of the cytosolic chaperonin complex, cause the misexpression of germline-specific genes in somatic tissues and this germline-like quality contributes to the increased survival and health of these animals (Fig. 5).

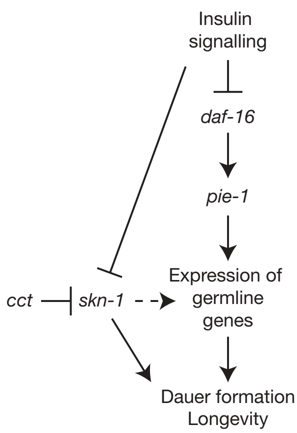


Figure 5 | Model for the regulation of germline gene expression in the soma by *C. elegans* longevity regulators. Insulin-like signalling negatively regulates the FOXO transcription factor DAF-16. DAF-16 positively regulates the expression of *pie-1* and germline-specific genes in somatic tissues. The cytosolic chaperonin complex (*cct*) negatively regulates the activity of SKN-1 and the somatic expression of PGL-1. Solid lines represent experimentally observed regulation, dashed lines indicate potential regulation.

Discussion

Our data show that many *C. elegans* longevity mutants exhibit a soma-to-germline transformation that contributes to their enhanced survival. The somatic protection stemming from the soma-to-germline transformation probably has a role during both the pre-reproductive dauer diapause stage, in which *C. elegans* can survive for months, and the post-reproductive adulthood phase, to promote longevity. This phenotype, although surprising, makes intuitive sense as germ cells use protective mechanisms to ensure genetic integrity. The recruitment of these normally germline-protective pathways in somatic tissues provides a mechanism for the finding that long-lived mutants exhibit resistance to genotoxic stress. Germ cells are tremendously resistant to numerous stresses, which renders them immortal and capable of regeneration on a scale not shared by somatic tissues. We propose that the transformation of somatic cells to a more germ-cell-like state provides increased genomic stability and may be an important pathway facilitating lifespan extension.

Alternatively, misregulation of germline genes in the soma may result in the ability of these post-mitotic cells to reinstate a proliferative state. *chk-1* and other checkpoint proteins regulate post-mitotic cell survival[38]. In the embryo the CHK-1 pathway is suppressed so that cell divisions can occur properly. If in the insulin-like signalling mutants the somatic tissues were more proliferative, this could also result in increased lifespan. Notably, *smk-1* (also known as *rad-2*) regulates the silencing of *chk-1*, and SMK-1 is a component of the insulin-like signalling longevity pathway[39,40]. A return to a replicatively competent state could also explain the observed stability of the somatic nuclear pore complex in the insulin signalling mutants if the subunits are being replaced with new copies, which does not occur in non-proliferative cell types[41]. This idea fits with the observation that RNAi towards the nuclear pore component *dct-3* (which encodes a homologue of human TPR) increases lifespan in wild-type animals but decreases the fitness of *daf-2* mutants[42], similar to our analysis of the germline-misexpressed genes.

The fact that the most potent transcriptional regulator of the lifespan extension phenotypes in insulin-like signalling mutants, *daf-16*, mediates the soma-to-germline transformation, points to the importance of this phenotype in the regulation of lifespan and stress resistance[14,15]. We demonstrate here by EMSA and ChIP that DAF-16 can bind to the *pie-1* promoter in somatic cells. The *in vivo* binding is weaker than the previously described target promoters of *sod-3* and *mtl-1*; however, this level of binding correlates with a lower induction of *pie-1* mRNA compared to the canonical targets (Supplementary Fig. 5). A similar level of enrichment of DAF-16 on the *pie-1* promoter in a strain expressing DAF-16::mCherry in the hypodermis, muscle and neurons, but not the intestine, points to these cell types as the key regulatory tissues (Supplementary Table 3)[43–45]. In conjunction with our qRT–PCR and genetic epistasis analyses, we have uncovered that DAF-16 is an important direct regulator of *pie-1* expression and modulator of the soma-to-germline transformation phenotype.

Our data support a model in which the somatic tissues of insulin-like signalling mutants display transcriptional dysregulation to become more germ-cell-like. Unlike the *cct* RNAi-treated animals, the detection of misexpressed germline proteins in the insulin-like signalling mutants has been complicated. However, germ cells are also under tight translational control, which may explain the lowered level of germline proteins[46]. This added layer of expression control might be important for the longevity phenotype, because canonical synMuvB mutants such as *mep-1*, *lin-15b* and *lin-35* strongly misexpress germline proteins in the soma[5] and have general sickness and a short lifespan (data not shown and Supplementary Fig. 11).

Similar to the longevity and soma-to-germline transformation, the enhanced RNAi phenotype of the insulin-like signalling mutants is dependent on DAF-16 (ref. 3). Furthermore, loss of the DAF-16 transcriptional target gene *zfp-1* (homologous to human *AF10*)[22] can suppress the multivulval phenotype of synMuv mutants and

partially suppress the lethality of *mep-1* mutants[34]. Moreover, synMuvB mutants transcriptionally dysregulate the gonad distal tip cell marker *lag-2* in other somatic tissues, which is partially dependent on *zfp-1*[34]. Finally, *zfp-1* has been shown to be required for the ability of *C. elegans* to perform RNAi[47]. Taken together, these data suggest that soma-to-germline transformation, enhanced RNAi, and longevity share common regulatory mechanisms.

The idea that somatic tissues have the potential of adopting a more germ-cell-like character provides a means of dissecting the dichotomy between the instability of somatic cells and the immortality of germline stem cells. Uncovering the mechanisms that mediate this transformation phenotype will provide insight into the pathways that maintain germ cell viability, cell fate specificity and increased survival stemming from the acquisition of germline quality by somatic cells.

Given that protection of the germ line is an evolutionarily shared trait across species, it will be interesting to investigate whether this is a broadly conserved mechanism of modulating lifespan. The idea that somatic cells maintain the potential to reacquire pathways lost during differentiation is tantalizing and may facilitate the elucidation of therapeutics to assist in cellular repair and possibly regeneration.

METHODS SUMMARY

C. elegans were cultured using standard techniques[48] and fed on *Escherichia coli* OP50 or HT115 containing a dsRNA-expressing plasmid[49].

Microscopy. Worms were paralysed with 2,3-butanedione monoxime, mounted on agarose pads, and imaged on a Zeiss Axioplan microscope. Live worms were imaged with a $\times 10$ objective under a Zeiss Discovery Microscope.

Quantitative RT–PCR. Synchronized animals were removed from growth plates and thoroughly washed in M9 buffer. RNA was isolated by Trizol (Invitrogen) and potential DNA contamination was removed by TurboDNA-Free (Ambion). The purity/stability of the RNA was tested by gel electrophoresis and ultraviolet spectroscopy. Complementary DNA was synthesized from 2 μg of RNA with SuperscriptIII reverse transcriptase (Invitrogen), and quantitative PCR was performed with SYBR Green (BioRad) on a BioRad iCycler. Three biological replicate samples were tested in triplicate. Samples were normalized to *rpl-32* and/or *snb-1* transcript levels.

Lifespan assays. Post-developmental RNAi lifespan analysis was performed as previously described[1]. In brief, synchronized worms fed on vector control bacteria were transferred to RNAi bacteria on day 1 of adulthood and traditional lifespan analysis was performed.

Statistical analysis. Statistical analyses were performed using the software SPSS (SPSS Inc.). The survival rate of each RNAi-treated population is compared with that of the population treated with control RNAi using the log-rank test. $P < 0.05$ was considered as significantly different from control.

Immunostaining. Animals were permeabilized using a freeze–crack method by freezing between a polylysine-coated slide and a coverslip, followed by rapid removal of the coverslip. Slides were immediately immersed in ice-cold 100% methanol for 120 min, followed by ice-cold 100% acetone for 15 min. Larvae were stained using monoclonal (1:40 dilution in PBSTB (PBS containing 0.1% Tween20 and 1% BSA)) anti-PGL-1 (K76) antibodies (gift from C. Mello) overnight at 4 °C, followed by Alexa Fluor-conjugated goat anti-mouse IgM (1:100 dilution in PBST; Invitrogen).

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 4 November 2008; accepted 1 May 2009.
Published online 7 June 2009.

1. Curran, S. P. & Ruvkun, G. Lifespan regulation by evolutionarily conserved genes essential for viability. *PLoS Genet.* **3**, e56 (2007).
2. Garsin, D. A. *et al.* Long-lived *C. elegans daf-2* mutants are resistant to bacterial pathogens. *Science* **300**, 1921 (2003).
3. Wang, D. & Ruvkun, G. Regulation of *Caenorhabditis elegans* RNA interference by the daf-2 insulin stress and longevity signaling pathway. *Cold Spring Harb. Symp. Quant. Biol.* **69**, 429–431 (2004).
4. Wilkins, C. *et al.* RNA interference is an antiviral defence mechanism in *Caenorhabditis elegans. Nature* **436**, 1044–1047 (2005).
5. Wang, D. *et al.* Somatic misexpression of germline P granules and enhanced RNA interference in retinoblastoma pathway mutants. *Nature* **436**, 593–597 (2005).
6. Batista, P. J. *et al.* PRG-1 and 21U-RNAs interact to form the piRNA complex required for fertility in *C. elegans. Mol. Cell* **31**, 67–78 (2008).
7. O'Neil, N. & Rose, A. DNA repair. *WormBook* **Jan 13**, 1–12 (2006).
8. Das, P. P. *et al.* Piwi and piRNAs act upstream of an endogenous siRNA pathway to suppress Tc3 transposon mobility in the *Caenorhabditis elegans* germline. *Mol. Cell* **31**, 79–90 (2008).
9. Gartner, A., Boag, P. R. & Blackwell, T. K. Germline survival and apoptosis. *WormBook* **Sept 14**, 1–20 (2008).
10. Robert, V. J., Sijen, T., van Wolfswinkel, J. & Plasterk, R. H. Chromatin and RNAi factors protect the *C. elegans* germline against repetitive sequences. *Genes Dev.* **19**, 782–787 (2005).
11. Kawasaki, I. *et al.* PGL-1, a predicted RNA-binding component of germ granules, is essential for fertility in *C. elegans. Cell* **94**, 635–645 (1998).
12. Mello, C. C., Draper, B., Krause, M., Weintraub, H. & Priess, J. The *pie-1* and *mex-1* genes and maternal control of blastomere identity in early *C. elegans* embryos. *Cell* **70**, 163–176 (1992).
13. Narbonne, P. & Roy, R. Inhibition of germline proliferation during *C. elegans* dauer development requires PTEN, LKB1 and AMPK signalling. *Development* **133**, 611–619 (2006).
14. Ogg, S. *et al.* The Fork head transcription factor DAF-16 transduces insulin-like metabolic and longevity signals in *C. elegans. Nature* **389**, 994–999 (1997).
15. Lin, K., Dorman, J. B., Rodan, A. & Kenyon, C. *daf-16*: An HNF-3/forkhead family member that can function to double the life-span of *Caenorhabditis elegans. Science* **278**, 1319–1322 (1997).
16. Kenyon, C., Chang, J., Gensch, E., Rudner, A. & Tabtiang, R. A. C. *elegans* mutant that lives twice as long as wild type. *Nature* **366**, 461–464 (1993).
17. Larsen, P. L. & Clarke, C. F. Extension of life-span in *Caenorhabditis elegans* by a diet lacking coenzyme Q. *Science* **295**, 120–123 (2002).
18. Houthoofd, K., Gems, D., Johnson, T. E. & Vanfleteren, J. R. Dietary restriction in the nematode *Caenorhabditis elegans. Interdiscip. Top. Gerontol.* **35**, 98–114 (2007).
19. Kaeberlein, T. L. *et al.* Lifespan extension in *Caenorhabditis elegans* by complete removal of food. *Aging Cell* **5**, 487–494 (2006).
20. Furuyama, T., Nakazawa, T., Nakano, I. & Mori, N. Identification of the differential distribution patterns of mRNAs and consensus binding sequences for mouse DAF-16 homologues. *Biochem. J.* **349**, 629–634 (2000).
21. Murphy, C. T. *et al.* Genes that act downstream of DAF-16 to influence the lifespan of *Caenorhabditis elegans. Nature* **424**, 277–283 (2003).
22. Oh, S. W. *et al.* Identification of direct DAF-16 targets controlling longevity, metabolism and diapause by chromatin immunoprecipitation. *Nature Genet.* **38**, 251–257 (2006).
23. Li, J. *et al. Caenorhabditis elegans* HCF-1 functions in longevity maintenance as a DAF-16 regulator. *PLoS Biol.* **6**, e233 (2008).
24. Dalley, B. K. & Golomb, M. Gene expression in the *Caenorhabditis elegans* dauer larva: developmental regulation of Hsp90 and other genes. *Dev. Biol.* **151**, 80–90 (1992).
25. Batchelder, C. *et al.* Transcriptional repression by the *Caenorhabditis elegans* germ-line protein PIE-1. *Genes Dev.* **13**, 202–212 (1999).
26. Unhavaithaya, Y. *et al.* MEP-1 and a homolog of the NURD complex component Mi-2 act together to maintain germline-soma distinctions in *C. elegans. Cell* **111**, 991–1002 (2002).
27. Beanan, M. J. & Strome, S. Characterization of a germ-line proliferation mutation in *C. elegans. Development* **116**, 755–766 (1992).
28. Arantes-Oliveira, N., Apfeld, J., Dillin, A. & Kenyon, C. Regulation of life-span by germ-line stem cells in *Caenorhabditis elegans. Science* **295**, 502–505 (2002).
29. Vijg, J. & Campisi, J. Puzzles, promises and a cure for ageing. *Nature* **454**, 1065–1071 (2008).
30. Kenyon, C. The plasticity of aging: insights from long-lived mutants. *Cell* **120**, 449–460 (2005).
31. van Haaften, G. *et al.* Identification of conserved pathways of DNA-damage response and radiation protection by genome-wide RNAi. *Curr. Biol.* **16**, 1344–1350 (2006).
32. Pothof, J. *et al.* Identification of genes that protect the *C. elegans* genome against mutations by genome-wide RNAi. *Genes Dev.* **17**, 443–448 (2003).
33. Hsin, H. & Kenyon, C. Signals from the reproductive system regulate the lifespan of *C. elegans. Nature* **399**, 362–366 (1999).
34. Cui, M., Kim, E. B. & Han, M. Diverse chromatin remodeling genes antagonize the Rb-involved SynMuv pathways in *C. elegans. PLoS Genet.* **2**, e74 (2006).
35. Kahn, N. W., Rea, S. L., Moyle, S., Kell, A. & Johnson, T. E. Proteasomal dysfunction activates the transcription factor SKN-1 and produces a selective oxidative-stress response in *Caenorhabditis elegans. Biochem. J.* **409**, 205–213 (2008).
36. An, J. H. & Blackwell, T. SKN-1 links *C. elegans* mesendodermal specification to a conserved oxidative stress response. *Genes Dev.* **17**, 1882–1893 (2003).
37. Tullet, J. M. *et al.* Direct inhibition of the longevity-promoting factor SKN-1 by insulin-like signaling in *C. elegans. Cell* **132**, 1025–1038 (2008).
38. Olsen, A., Vantipalli, M. C. & Lithgow, G. J. Checkpoint proteins control survival of the postmitotic cells in *Caenorhabditis elegans. Science* **312**, 1381–1385 (2006).
39. Kim, S. H., Holway, A. H., Wolff, S., Dillin, A. & Michael, W. M. SMK-1/PPH-4.1-mediated silencing of the CHK-1 response to DNA damage in early *C. elegans* embryos. *J. Cell Biol.* **179**, 41–52 (2007).
40. Wolff, S. *et al.* SMK-1, an essential regulator of DAF-16-mediated longevity. *Cell* **124**, 1039–1053 (2006).
41. D'Angelo, M. A., Raices, M., Panowski, S. H. & Hetzer, M. W. Age-dependent deterioration of nuclear pore complexes causes a loss of nuclear integrity in postmitotic cells. *Cell* **136**, 284–295 (2009).
42. Pinkston-Gosse, J. & Kenyon, C. DAF-16/FOXO targets genes that regulate tumor growth in *Caenorhabditis elegans. Nature Genet.* **39**, 1403–1409 (2007).

43. Libina, N., Berman, J. R. & Kenyon, C. Tissue-specific activities of *C. elegans* DAF-16 in the regulation of lifespan. *Cell* **115,** 489–502 (2003).
44. Wolkow, C. A., Kimura, K. D., Lee, M. S. & Ruvkun, G. Regulation of *C. elegans* life-span by insulinlike signaling in the nervous system. *Science* **290,** 147–150 (2000).
45. Murphy, C. T., Lee, S. J. & Kenyon, C. Tissue entrainment by feedback regulation of insulin gene expression in the endoderm of *Caenorhabditis elegans. Proc. Natl Acad. Sci. USA* **104,** 19046–19050 (2007).
46. Lee, M. H. & Schedl, T. Identification of *in vivo* mRNA targets of GLD-1, a maxi-KH motif containing protein required for *C. elegans* germ cell development. *Genes Dev.* **15,** 2408–2420 (2001).
47. Kim, J. K. *et al.* Functional genomic analysis of RNA interference in *C. elegans. Science* **308,** 1164–1167 (2005).
48. Brenner, S. The genetics of *Caenorhabditis elegans. Genetics* **77,** 71–94 (1974).
49. Fire, A. *et al.* Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans. Nature* **391,** 806–811 (1998).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements Some strains were provided by the *Caenorhabditis* Genetics Center, which is funded by the National Institutes of Health National Center for Research Resources. We thank A. Conery, H. Gabel, J. Melo, E. O'Rourke, D. Simon, A. Soukas and M. Wang for reagents, discussion and reading of the manuscript; G. Seydoux and members of the Ruvkun, Ausubel and Kaplan laboratories for discussion; S. Strome for the *pie-1p::gfp::pgl-1* strain and C. Mello for monoclonal antisera against PGL-1, S. S. Lee, S. Ercan and the Kingston laboratory for technical advice on ChIP. This work was supported by grants from the National Institutes of Health and the National Institute on Aging F32-AG026207 to S.P.C. and R01-AG016636 to G.R. and the European Molecular Biology Organization and the Human Frontier Science Program to C.G.R.

Author Contributions S.P.C. and G.R. conceived and designed the experiments. S.P.C., X.W. and C.G.R. performed the experiments. S.P.C. and G.R. wrote the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to G.R. (ruvkun@molbio.mgh.harvard.edu).

METHODS

Strains and alleles. Strains were maintained and cultured using standard techniques. Wild-type N2 Bristol, GR1373 *eri-1(mg366)IV*, DR1572 *daf-2(e1368)III*, CB1370 *daf-2(e1370)III*, GR1309 *daf-16(mgDf47)I; daf-2(e1370)III*, *age-1(mg305)*, SS747 *Is[pie-1p::gfp::pgl-1]*, *daf-2(e1370); Is[pie-1p::gfp::pgl-1]*, *age-1(mg305); Is[pie-1p::gfp::pgl-1]*, NL3400 *pkIs1604[hsp-16.2p::ATG(A_{17})gfp::lacZ; pRF4 (rol-6(su1006))]*, *daf-2(e1368); pkIs1604[hsp-16.2p::ATG(A_{17})gfp::lacZ; pRF4 (rol-6(su1006))]*, TJ356 *[daf-16p::daf-16::gfp (rol-6(su1006))]*.

DAF-16–Flag purification. *daf-16a* cDNA was amplified using a 3′ primer encoding a Flag epitope tag and subsequently cloned into vector pEXP5-NT (Invitrogen) downstream of a 6×His coding sequence. This vector was cloned into BL21(DE3) cells and protein expression was induced with isopropyl-β-D-thiogalactoside. 6×His–DAF-16–Flag was purified over NiNTA resin (Qiagen) under native conditions and eluted with imidazol. Purification was monitored by SDS–PAGE and Coomassie staining. The purified protein sample was dialysed against PBS and subsequently used for EMSA.

Electrophoretic mobility shift assays. Fragments of the *pie-1* promoter were amplified from *C. elegans* genomic DNA by PCR and end-labelled with ^{32}P. The radiolabelled fragments were incubated with the purified recombinant 6×His–DAF-16–Flag protein in binding buffer (16% Ficoll, 40 ng μl^{-1} sheared salmon sperm DNA, 1× TE) on ice. One-third of this reaction was incubated with a cold dsDNA oligonucleotide containing several DAF-16-associated/binding elements, one-third of the reaction was incubated with anti-Flag antibody (Sigma), and one-third was left untreated. The reactions were stopped with OrangeG glycerol DNA loading buffer and resolved on a 4% acrylamide/TBE gel at 4 °C. Gels were dried on Whatman paper and visualized by autoradiography.

Western blot analysis. Synchronized populations of worms were removed from growth plates with M9 buffer, washed three times with PBS buffer, and frozen in liquid nitrogen. Samples were lysed by sonication. Protein concentration was determined by BCA assays (Pierce). Equal amounts of protein were denatured in SDS sample buffer at 95 °C then resolved on NuPage 4–12% SDS acrylamide gels (Invitrogen) and subsequently transferred to PVDF membranes, blocked with milk and probed with antisera specific for GFP (Clontech) or actin (MBP) and visualized by SuperSignal ECL reagent (Pierce).

EMSA oligonucleotides. DAF-16-associated element forward (F), 5′-GTAAACAACTTATCAGTAAATAACTTATCAGTAAACAACTTATCAGTAAATAACTTATCA-3′; DAF-16-associated element reverse (R), 5′-TGATAAGTTATTTACTGATAAGTTGTTTACTGATAAGTTATTTACTGATAAGTTGTTTAC-3′; *pie-1p*(-1778F), 5′-CGCTACGAGATACTTTGCGCGCCAAATATG-3′; *pie-1p* (-1571R), 5′-CGCTACGAGATACTTTGCGCGCCAAATATG-3′; *pie-1p* (-1278F), 5′-GCAAGCGCGCTCCAACCAACGAGTTCAA-3′; *pie-1p* (-1120R), 5′-CATGCCTATGGGGACACAATGACCG-3′; *pie-1p* (-983F), 5′-CTTTACGCCACATAAAGTGATATGGAAGAG-3′; *pie-1p* (-708R), 5′-CGCTGAAGGTCATTTCTGTTGTCCTCACC-3′; *pie-1p* (-736F), 5′-GGTGAGGACAACAGAAATGACCTTC-3′; *pie-1p* (-491R), 5′-GGCGCATGCCTCAAGTTGTGCTTTCGTTGCG-3′; *pie-1p* (-521F), 5′-CGCAACGAAAGCACAACTTGAGGCA-3′; *pie-1p* (-295R), 5′-GAGAAGTGCTCCGTGCTGAATTTGGG-3′; *pie-1p* (-308F), 5′-CCCAAATTCAGCACGGAGCACTTCTC-3′; *pie-1p* (ATGR), 5′-CGGCAATCGGCTTTGTTTGAGCCAT-3′.

Quantitative RT–PCR oligonucleotides. Primers used were *pie-1* qPCR, 5′-AGCGCTCAAGATCACTCCAT-3′, 5′-GTGCAATTGTCGTTGTACGG-3′; *pgl-1* qPCR, 5′-CGCTACTCCAATGCCTCTTC-3′, 5′-ACTGGTGGGAGCAAATTGAC-3′; *pgl-2* qPCR, 5′-GGGATCCTCGACTCACTGAA-3′, 5′-TCCAAATTGCGCATGAAATA-3′; *pgl-3* qPCR, 5′-TTGCAACGGAGTGTTTTCAA-3′, 5′-GCTCTTCTTCTCCACGTGCT-3′; *rpl-32* qPCR, 5′-AGGGAATTGATAACCGTGTCCGCA-3′, T5′-GTAGGACTGCATGAGGAGCATGT-3′; *snb-1* qPCR, 5′-CCGGATAAGACCATCTTGACG-3′, 5′-GACGACTTCATCAACCTGAGC-3′.

Chromatin immunoprecipitation. ChIP was performed as previously described[23] (and S. Ercan, personal communication; S. Lee, personal communication). In brief, ground frozen worm powder was crosslinked using 1% formaldehyde in PBS and subjected to sonication. Immunoprecipitation was performed using Protein A Dynabeads (Invitrogen) and polyclonal rabbit-anti-GFP antibody (Clontech) or polyclonal rabbit-anti-mCherry antibody. Protein–DNA complexes were then eluted from protein A beads and treated with RNase A and proteinase K. Precipitated DNA fragments were purified and subjected to qPCR analysis.

qPCR primers for ChIP. Primers for the non-coding chromosomal region Chr IV and for the promoter of *mtl-1* were previously described[23].

ChIP primers. Non-coding region Chr V, 5′-ACTGGGAGTGGTTCAAGTGG-3′, 5′-TACCTCCAGCACCACCTTCT-3′; *sod-3*, 5′-ACAACAATGTGCTGGCCTTG-3′, 5′-AATGCATTTCGGGACGTTAG-3′; *pie-1* A, 5′-TCCCGCAACGAAAATAAAAC-3′, 5′-AAAAACGGCGAGAAATCCAT-3′; *pie-1* B, 5′-CAAAAATCGAGAAAAATGCTTG-3′, 5′-TTGTGCTTTCGTTGCGATAC-3′.

Vol 459 | 25 June 2009 | doi:10.1038/nature08109

GOLPH3 modulates mTOR signalling and rapamycin sensitivity in cancer

Kenneth L. Scott[1]*, Omar Kabbarah[1]*, Mei-Chih Liang[1,4], Elena Ivanova[2], Valsamo Anagnostou[5], Joyce Wu[1], Sabin Dhakal[1], Min Wu[1], Shujuan Chen[1], Tamar Feinberg[1], Joseph Huang[1], Abdel Saci[6], Hans R. Widlund[3,7], David E. Fisher[3,8], Yonghong Xiao[2], David L. Rimm[5], Alexei Protopopov[2], Kwok-Kin Wong[1,4] & Lynda Chin[1,2,7]

Genome-wide copy number analyses of human cancers identified a frequent 5p13 amplification in several solid tumour types, including lung (56%), ovarian (38%), breast (32%), prostate (37%) and melanoma (32%). Here, using integrative analysis of a genomic profile of the region, we identify a Golgi protein, GOLPH3, as a candidate targeted for amplification. Gain- and loss-of-function studies *in vitro* and *in vivo* validated *GOLPH3* as a potent oncogene. Physically, GOLPH3 localizes to the *trans*-Golgi network and interacts with components of the retromer complex, which in yeast has been linked to target of rapamycin (TOR) signalling. Mechanistically, GOLPH3 regulates cell size, enhances growth-factor-induced mTOR (also known as FRAP1) signalling in human cancer cells, and alters the response to an mTOR inhibitor *in vivo*. Thus, genomic and genetic, biological, functional and biochemical data in yeast and humans establishes *GOLPH3* as a new oncogene that is commonly targeted for amplification in human cancer, and is capable of modulating the response to rapamycin, a cancer drug in clinical use.

Cellular growth, proliferation and survival are regulated by a complex network of intracellular and extracellular signal transduction cascades. The growth-factor-responsive receptor tyrosine kinase (RTK)–phosphatidylinositol-3-kinase (PI3K) pathway has an important role in governing these processes[1]. The serine/threonine kinase AKT functions as a central integrator of RTK–PI3K signalling to modulate downstream effectors, most notably the TSC1/2–mTOR complexes. mTOR, a serine/threonine protein kinase and 'target of rapamycin', acts as a primary regulator of protein synthesis and cell growth[2]. Genetic studies in *Drosophila* and mice[3–6] have shown that mTOR activity can influence cell size, an important parameter governing entry into the cell cycle[7]. mTOR also integrates diverse upstream signals that include amino acid and energy stress sensing to regulate cell proliferation, growth and survival[8,9]. mTOR is present in two separate signalling complexes, mTOR complex 1 (mTORC1) and mTORC2, which differ in their subunit composition and their sensitivity to the bacterial macrolide rapamycin. Rapamycin inhibits mTOR activity when bound to the protein raptor, leading to reduced cell growth, cell size and proliferation[2,10,11].

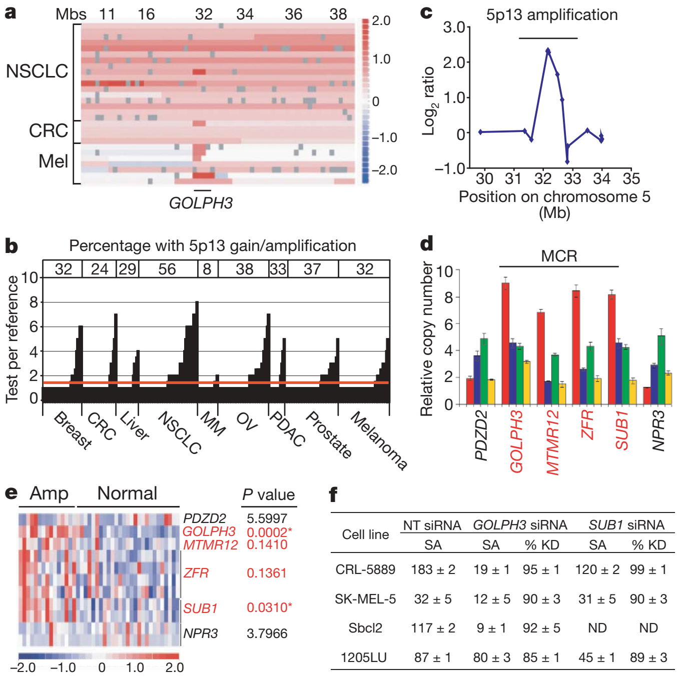
A hallmark feature of human cancer is its highly rearranged genome, manifesting as numerous copy-number amplifications and deletions detectable by genome-wide array-based comparative genome hybridization (array-CGH) profiling. Reasoning that genomic alterations recurrent in different human cancer types are more likely to represent 'driver' events targeting cancer genes of broad biological and pathogenetic significance, we defined a focal amplification within a larger 5p13 regional gain present at high frequencies in several solid tumours, including melanoma, colon adenocarcinoma and non-small-cell lung cancer (NSCLC). Integrative genomics combined with clinicopathological and functional validation identified a Golgi-localizing protein, GOLPH3, as a new oncoprotein targeted for gene amplification at 5p13. Mechanistic studies using both yeast and human systems linked GOLPH3 to the retromer protein trafficking complex, mTOR signalling and the response to rapamycin.

GOLPH3 is a target of 5p13 amplification

Array-CGH analysis of 83 melanoma specimens showed a focal amplification within a larger 5p13 regional copy number gain, which was also present in NSCLC and colon adenocarcinomas (Fig. 1a), prompting a broad survey by fluorescence *in situ* hybridization (FISH). Analysis on tumour tissue microarrays (TMA) containing 307 cores of diverse tumour types showed that 5p13 gain was present significantly in all tumour types surveyed, including 56% (27 out of 48) of NSCLC cores, 38% (18 out of 48) of ovarian carcinoma cores, 37% (16 out of 43) of prostate cancer cores, as well as 32% (12 out of 38) of melanoma cores (Fig. 1b, Supplementary Fig. 1 and Supplementary Table 1). Quantitative real-time PCR across the 5p13 region in informative tumour samples delimited a 0.8 megabase (Mb) minimal common region (MCR) encompassing four resident annotated genes (Fig. 1c, d). Given that copy-number aberration (CNA) is a mechanism to drive deregulated gene expression, we next investigated the expression pattern of these resident genes in an NSCLC collection with matched expression and array-CGH profiles. As shown in Fig. 1e, only *GOLPH3* and *SUB1*, but not *MTMR12* and *ZFR*, showed statistically significant correlation between expression level and copy-number status, thereby pointing to *GOLPH3* and *SUB1* as viable candidate target(s) of this amplification.

To assess the cancer-relevance of GOLPH3, SUB1 or both, knockdown assays using pooled short interfering RNAs (siRNAs) (Fig. 1f) were performed to gauge the dependence of the human tumour (NSCLC and melanoma) cell lines on either protein for their transformed phenotype relative to the underlying copy-number status

[1]Department of Medical Oncology, [2]Belfer Institute for Applied Cancer Science, [3]Department of Pediatric Oncology, Dana-Farber Cancer Institute, [4]Ludwig Center at Dana-Farber/Harvard Cancer Center, Boston, Massachusetts 02115, USA. [5]Department of Pathology, Yale University School of Medicine, New Haven, Connecticut 06520, USA. [6]Department of Systems Biology, Harvard Medical School, [7]Department of Dermatology, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts 02115, USA. [8]Department of Dermatology, Massachusetts General Hospital, Boston, Massachusetts 02114, USA.
*These authors contributed equally to this work.



Cell line	NT siRNA SA	GOLPH3 siRNA SA	GOLPH3 siRNA % KD	SUB1 siRNA SA	SUB1 siRNA % KD
CRL-5889	183 ± 2	19 ± 1	95 ± 1	120 ± 2	99 ± 1
SK-MEL-5	32 ± 5	12 ± 5	90 ± 3	31 ± 5	90 ± 3
Sbcl2	117 ± 2	9 ± 1	92 ± 5	ND	ND
1205LU	87 ± 1	80 ± 3	85 ± 1	45 ± 1	89 ± 3

Figure 1 | Genomic characterization of 5p13 amplification. a, Array-CGH heat map detailing *GOLPH3* amplification at 5p13 in representative tumour specimens and cell lines from malignant melanoma (Mel), colon adenocarcinoma (CRC) and non-small-cell lung cancer (NSCLC). Regions of genomic amplification and deletion are denoted in red and blue, respectively. Mbs denote the position on chromosome 5 in Mb. **b**, Histogram summary of copy-number status at 5p13 by TMA-FISH analysis of 307 tumour cores of the indicated tumour types. MM, multiple myeloma; OV, ovarian carcinoma; PDAC, pancreatic ductal adenocarcinoma. For details see Supplementary Table 1. **c**, Minimum common region of the 5p13 amplicon defined by array-CGH from one representative tumour (melanoma C27) with focal amplification. The *y* axis is the $\log_2$ ratio compared to reference sample (pooled normal human DNA), and the *x* axis denotes the position on chromosome 5. **d**, Delimitation of chromosome 5p13 amplicon boundaries by genomic quantitative PCR using four informative cell line and tumour specimens. Samples include tumour specimens C27 (primary melanoma; red), C1 (primary melanoma; blue), CRL-5889 (NSCLC cell line; green) and Sbcl2 (melanoma cell line; yellow). Genes denoted by red font comprise 5p13 MCR; for details see Supplementary Information. Error bars indicate ± s.d. **e**, Heat map depiction of Affymetrix expression data for NSCLC 5p13-amplified (Amp) and 5p13-normal specimens. Asterisks denote significant correlation after Bonferroni correction for multiple testing. **f**, Summary of soft agar colony counts (SA) and corresponding siRNA knockdown (% KD) of GOLPH3 and SUB1 in the indicated cell lines with amplified or normal *GOLPH3* copy number. ND, not determined; NT, non-targeting.

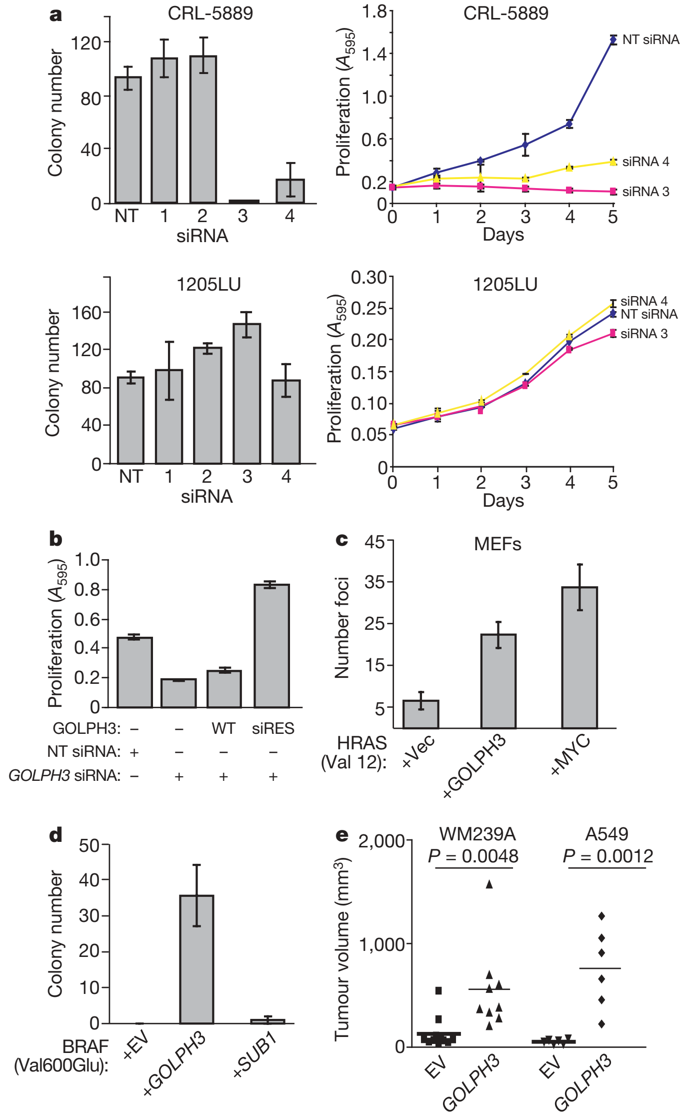


Figure 2 | Functional validation of GOLPH3. a, The indicated cell lines were treated with non-targeting (NT) siRNA or individual siRNAs against *GOLPH3* (siRNA 1–4) to assay for effects on anchorage-independent growth in soft agar (left) and cell proliferation (right). Error bars indicate ± s.d. **b**, A549 parental cells and those expressing either wild type (WT) or siRNA-resistant *GOLPH3* (siRES) were treated with either non-targeting siRNA or siRNA against *GOLPH3*, and assessed for effects on cell proliferation. Error bars indicate ± s.d. Shown are the end-point values for day 5. **c**, Primary *Ink4a/Arf*-deficient MEFs were transfected with the indicated vectors expressing HRAS(Val 12), MYC and GOLPH3. Vec denotes LacZ vector control; error bars indicate ± s.d. Two-tailed *t*-test: HRAS(Val 12) + GOLPH3 versus HRAS(Val 12) + Vec, $P = 0.0018$. **d**, TERT-immortalized human melanocytes expressing activated BRAF(Val600Glu) were transduced with either *GOLPH3* or *SUB1* to assay for the effect on anchorage-independent growth in soft agar. Error bars indicate ± s.d. Two-tailed *t*-test for colony number: empty vector (EV) versus *GOLPH3*, $P = 0.0020$; empty vector versus *SUB1*, $P = 0.3739$. **e**, The indicated cell lines were transduced with *GOLPH3* to assay for the effect on growth of mouse xenograft tumours.

and overall expression level (Supplementary Fig. 2a). Knockdown of GOLPH3 resulted in considerable loss of anchorage-independent growth in CRL-5889 (NSCLC), Sbcl2 and SK-MEL-5 (melanoma) cells—three human cancer cell lines with 5p13 amplification and high expression level. However, a similar level of knockdown in 1205LU—a melanoma cell line without the 5p13 CNA and with low protein expression—resulted in minimal effect on anchorage independence (Fig. 1f). In contrast, equally effective knockdown of SUB1 in the 5p13-amplified tumour lines had either no or relatively modest effects on anchorage independence.

To confirm that the observed knockdown activity was not due to an off-target effect of the *GOLPH3* siRNA, we deconvoluted the siRNA pool and verified that two of the four independent siRNA duplexes (siRNAs 3 and 4) were effective at knocking down GOLPH3 (Supplementary Fig. 2b), which led to potent suppression of soft-agar growth and inhibition of proliferation in 5p13-amplified CRL-5889 cells (Fig. 2a). Similarly effective knockdown in 1205LU without 5p13 gain and low GOLPH3 expression showed minimal effect, indicating that acute GOLPH3 depletion is not generally toxic to all cells (Fig. 2a and Supplementary Fig. 2b). Notably, the specificity of siRNA 3 against *GOLPH3* was further documented by proliferation rescue of the NSCLC cell line, A549, by a *GOLPH3* complementary DNA engineered to be insensitive to siRNA 3 (siRES) (Fig. 2b). Together, these genetic loss-of-function studies using RNA interference (RNAi)-mediated knockdown pointed to GOLPH3 as the probable functionally active target of this amplification.

We next examined the impact of ectopic GOLPH3 expression in several model systems. GOLPH3 was capable of effecting malignant transformation of both primary non-transformed mouse and human cells. Specifically, in the classical co-transformation assay, GOLPH3 cooperated with activated HRAS(Val 12)—a constitutively active mutant in which Gly 12 is mutated to Val—to increase transformed focus formation in *Ink4a/Arf* (also known as *Cdkn2a*)-deficient primary mouse embryonic fibroblasts (MEFs) (Fig. 2c; 3.4-fold increase relative to HRAS(Val 12) alone). In primary human cells, GOLPH3 cooperated with oncogenic BRAF(Val600Glu) in TERT-immortalized human melanocytes (hereafter referred to as 'HMEL')[12] to confer anchorage-independent growth in soft agar, whereas SUB1 showed no transforming activity in this system (Fig. 2d). Similar activity was also observed in the 1205LU melanoma cell line (no 5p13 amplification, low GOLPH3 expression), in which GOLPH3 overexpression (Supplementary Fig. 2c) enhanced anchorage-independent growth and cell proliferation *in vitro* (Supplementary Fig. 2d and data not shown). Furthermore, GOLPH3 overexpression (Supplementary Fig. 2e, f) significantly enhanced xenotransplanted tumour growth (Fig. 2e) of human melanoma (WM239A) and NSCLC (A549) cell lines, both without 5p13 amplification. This series of reinforcing knockdown and overexpression studies demonstrates that *GOLPH3* is a bona fide oncogene with potent transforming activity.

Yeast genetics links GOLPH3 to mTOR

GOLPH3 (also known as GPP34, and GMx33 in rat) was initially identified as a peripheral membrane protein localized to the *trans*-Golgi network (TGN)[13,14]. Subsequent work with the rat homologue GMx33 showed that the protein is dynamically associated with the *trans*-Golgi matrix, rapidly moving from the TGN to the cytosol with localization in endosomes and at the plasma membrane[15]. As a class, TGN-localizing proteins have not been directly implicated in cancer pathogenesis. Thus, we first confirmed by confocal microscopy that both exogenously expressed and endogenous human GOLPH3 indeed colocalized to the Golgi apparatus via co-immunofluorescence with the TGN marker, TGN46, and to endosome-like structures as described previously[15] (Fig. 3a).

To gain mechanistic insights into the biological functions of GOLPH3, we screened for GOLPH3-interacting proteins using the yeast two-hybrid system (Supplementary Information and Supplementary Table 2). Most notable among the GOLPH3-interacting proteins was VPS35, a highly conserved member of the cargo-recognition complex of the retromer, which regulates retrograde transport of proteins that include transmembrane receptors from endosomes to the TGN[16]. After documenting a bait-dependent interaction of GOLPH3 with VPS35 in yeast (Supplementary Fig. 3a, b), we showed a physical interaction of VPS35 with both exogenously expressed and endogenous GOLPH3 in human cells by co-immunoprecipitation (Fig. 3b and Supplementary Fig. 3c). Confocal co-immunofluorescence studies further confirmed the colocalization of endogenous VPS35 and GOLPH3 at endosome-like structures (Fig. 3c).

A large-scale chemical genomic-profiling screen in *Saccharomyces cerevisiae*[17] has found that deletion mutants of *VPS35* and *VPS29* had altered sensitivity to rapamycin, an inhibitor of TOR signalling, suggesting that the retromer complex might function in the TOR signalling pathway in budding yeast. Thus, we postulated that GOLPH3 might regulate the mammalian orthologue of TOR (mTOR), thereby contributing to the pro-tumorigenic effects of GOLPH3. This hypothesis is supported by the observation that in human NSCLC tumour specimens, high 5p13 copy number was associated with increased mTOR expression and increased phosphorylation of the mTOR substrate, S6 kinase (S6K, also known as RPS6KB1) by automated

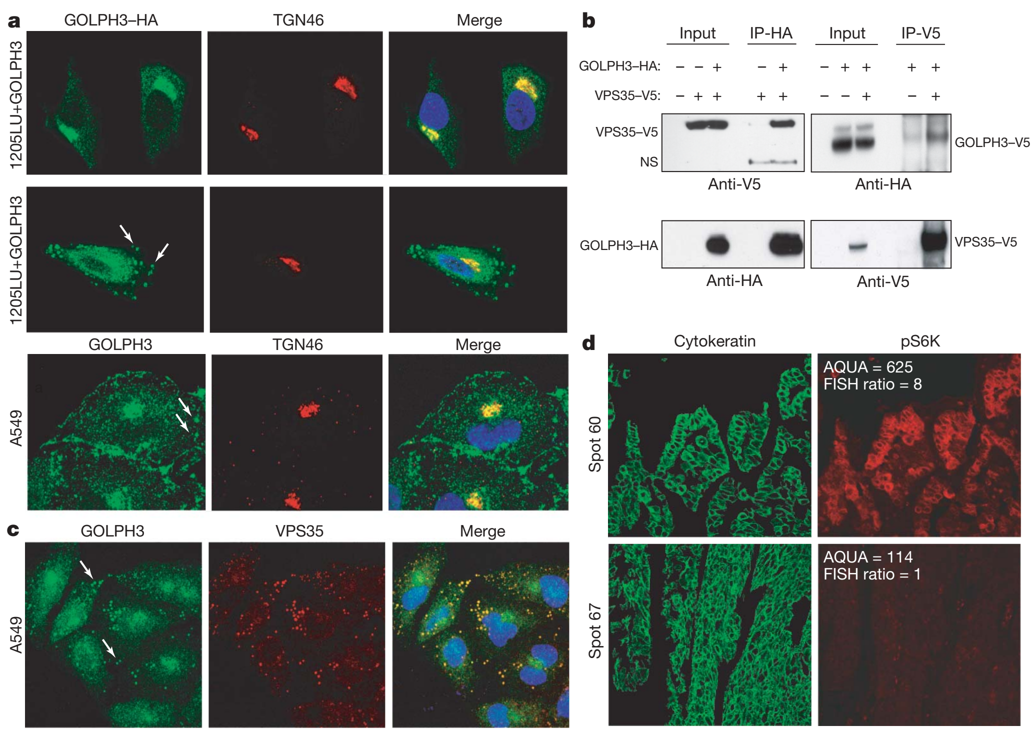


Figure 3 | GOLPH3 interacts with VPS35 and correlates with S6K phosphorylation. a, Top two panels, 1205LU melanoma cells stably expressing GOLPH3 were co-immunostained for haemagglutinin (HA) (GOLPH3–HA; green) and TGN46 (Golgi marker; red). Bottom panel, A549 cells were co-immunostained for GOLPH3 (green) and TGN46 (red). DNA was labelled with DAPI (4,6-diamidino-2-phenylindole; blue). Arrows indicate GOLPH3-positive endosome-like structures. **b,** The indicated constructs were transiently expressed in 293T cells, and proteins were immunoprecipitated (IP) with anti-HA (left) or anti-V5 (right) for immunoblotting with the indicated antibodies. NS, nonspecific band. **c,** A549 cells were co-immunostained for GOLPH3 (green) and VPS35 (red). DNA was labelled with DAPI (blue). Arrows indicate GOLPH3-positive co-staining at endosome structures. **d,** Automated quantitative analysis (AQUA) of pS6K(Thr 389) (red) in two representative lung adenocarcinomas. Cytokeratin (green) defines tumour and non-nuclear compartments. FISH ratio represents the 5p13:reference ratio as determined by FISH on consecutive TMA sections. Original magnification for AQUA images, ×20. Original magnification for **a** and **c**, ×100.

quantitative analysis (AQUA) immunofluorescence[18] (Fig. 3d and Supplementary Tables 3 and 4). Specifically, the mTOR expression level was associated with the cytoplasmic, but not nuclear nor total, phosphorylated-S6K(Thr 389) (pS6K) level ($r = 0.42$, $P = 0.001$, Pearson's correlation test). When the 5p13 copy number status was binarized into normal (with FISH-determined signal < 1.5) and gained (FISH signal ≥ 1.5), a significant correlation with increased mTOR (Spearman's correlation coefficient $= 0.475$, $P = 0.04$) and cytoplasmic pS6K (Spearman's correlation coefficient $= 0.724$, $P < 0.0001$) was observed, indicating that 5p13 CNA is positively correlated with mTOR–pS6K activity in the adenocarcinoma subtype of NSCLC. Notably, even when the 5p13 copy number was treated as a continuous variable, significant correlation was still observed with cytoplasmic pS6K in this subtype of NSCLC ($r = 0.513$, $P < 0.025$, Pearson's correlation test; Supplementary Table 4). Taken together, this correlative relationship in human cancers, coupled with the yeast genetic interaction data, supports the hypothesis that GOLPH3 regulates mTOR activity in mammalian cells.

GOLPH3 activates mTOR signalling

To test the hypothesis that GOLPH3 activates mTOR signalling, we first examined the biological consequences of GOLPH3 modulation. Consistent with the role of mTOR role in cell size regulation, RNAi-mediated GOLPH3 depletion led to a significant cell size reduction in A549 cells, an effect that was comparable to treatment with rapamycin[19] (Fig. 4a). Next, we assayed the biochemical consequences of GOLPH3 modulation.

Because the mTOR substrate S6K is a kinase effector of cell size that is phosphorylated at Thr 389 by mTOR[20,21], we investigated phosphorylated-S6K status as a readout of the mTORC1 axis. Consistent with the human tumour data showing increased pS6K levels in 5p13-amplified NSCLC specimens, GOLPH3 overexpression resulted in increased levels of pS6K in tumour cell lines (1205LU and A549), as well as in HMEL-tet-GOLPH3 cells—a TERT-immortalized human melanocyte cell line engineered with a tetracycline (tet)-regulated *GOLPH3* expression construct (Fig. 4b). Substantiating these observations using the inducible system, GOLPH3 induction further enhanced pS6K accumulation in response to growth-factor stimulation by epidermal growth factor (EGF; Fig. 4c). At the same time, we monitored phosphorylation of AKT (pAKT) at Ser 473, a direct substrate of mTORC2 (refs 22, 23). Similar to the mTORC1-mediated phosphorylation of S6K, we observed a comparable increase in pAKT phosphorylation in GOLPH3-overexpressing cells (Fig. 4b), suggesting that GOLPH3 can enhance signalling through both mTOR-associated complexes. Moreover, phosphorylation of AKT and S6K was significantly abrogated in *GOLPH3*-siRNA-treated NSCLC A549 and CRL-5889 cells compared to control cells in response to EGF (Fig. 4d, e). Further biochemical analyses showed altered phosphorylation of the mTOR substrates S6K(Thr 389), p4E-BP1(Thr 37/46) (also known as EIF4EBP1(Thr 37/46)) and AKT(Ser 473), with little to no affect on other signalling proteins including PTEN, MEK1/2 (MAP2K1/2), p44 (MAPK3 and ERK1) and P42MAPK (MAPK1 and ERK2), among others (Supplementary Fig. 4). Collectively, these data provided strong biochemical evidence that GOLPH3 activates mTOR signalling through phosphorylation of both mTORC1- and mTORC2-specific substrates.

GOLPH3 modulates rapamycin sensitivity

Complementing the genetic studies above, we next addressed whether GOLPH3 expression levels affect tumour cell sensitivity to pharmacological mTOR inhibition *in vivo*. Here we selected two human melanoma cell lines, 1205LU and WM239A, on the basis of their normal *GOLPH3* copy number and low protein expression, as well as their ability to readily form subcutaneous tumours *in vivo*. Parental cells were engineered to stably express either empty vector or GOLPH3, for orthotopic subcutaneous transplantation into immunodeficient animals for tumour growth. Next, we compared the degree of tumour growth inhibition in GOLPH3-expressing versus empty-vector control tumours on rapamycin treatment.

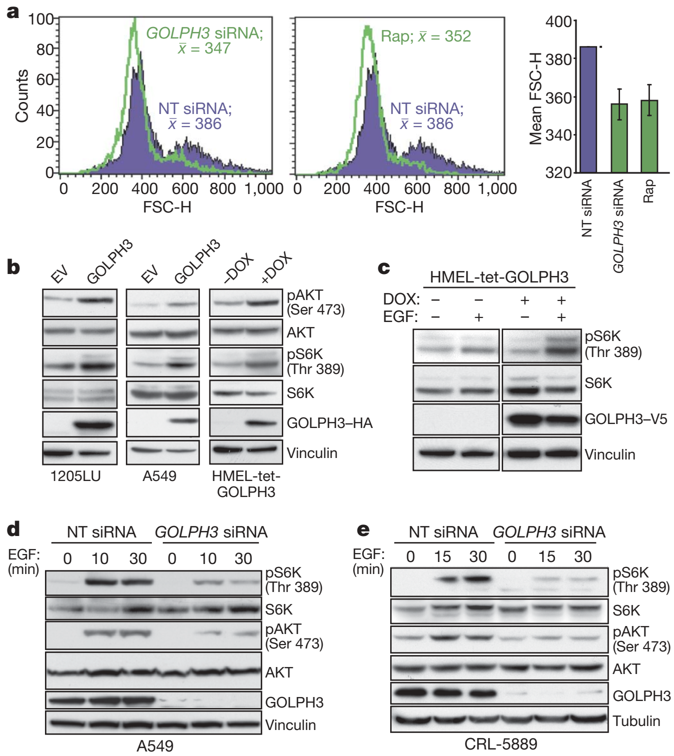
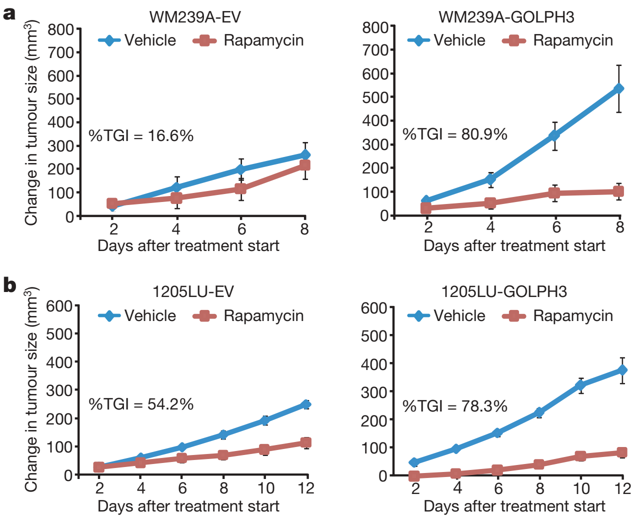


Figure 4 | GOLPH3 modulates cell size and phosphorylation status of mTOR substrates. **a**, Representative flow histograms for A549 cells treated with non-targeting (NT) siRNA (blue), siRNA against *GOLPH3* (left panel, green) or rapamycin (Rap; middle panel, green). Peak forward scatter-height (FSC-H) is indicated in histograms. The far right panel shows the mean FSC-H for several experiments ($n = 3$); error bars indicate ± s.d. **b**, Protein lysates extracted from 1205LU (left), A549 (middle) and HMEL-tet-GOLPH3 (right; with or without doxycycline (DOX)) cells expressing GOLPH3 were immunoblotted with the indicated antibodies. **c**, HMEL-tet-GOLPH3 cells were serum depleted and propagated with or without doxycycline, followed by treatment with or without EGF for 30 min for immunoblot analysis with the indicated antibodies. **d**, **e**, A549 (**d**) and CRL-5889 (**e**) cells were serum depleted and treated with either non-targeting or *GOLPH3* siRNA, followed by growth-factor stimulation with EGF for immunoblot analysis with the indicated antibodies.

Consistent with described earlier (Fig. 2e), 1205LU-GOLPH3 cells showed a significant growth advantage compared to 1205LU-empty-vector control cells *in vivo* (1.9-fold increase in tumour volume at 36 days post-injection in vehicle control cohort, $P = 0.0148$). After tumours reached a baseline volume of ~100 mm^3, the animals were randomized into control and treatment cohorts for intraperitoneal injection of either vehicle or rapamycin (6.0 mg kg^{-1}) every other day. The treatment trial was terminated when one animal in any cohort had to be euthanized for tumour burden according to Institutional Animal Care and Use Committee regulations. The inhibitory effect of rapamycin on mTOR activity of treated tumours was verified by western analysis (Supplementary Fig. 2g). The efficacy of rapamycin treatment was then calculated as the percentage of tumour growth inhibition of treated versus non-treated cohorts, after four doses (day 8 of trial) for WM239A cells, and six doses (day 12 of trial) for 1205LU cells. Indeed, GOLPH3-expressing tumours were significantly more sensitive to rapamycin *in vivo* (Fig. 5a–c and Supplementary Fig. 5). Therefore, we conclude that the biochemical effect of GOLPH3 on mTOR signalling is a critical aspect of its oncogenic function, because inhibition by rapamycin effectively blocked the growth advantage conferred by GOLPH3 *in vivo*.



c

Treatment		WM239A		1205LU	
		Size (mm³)	Fold change	Size (mm³)	Fold change
EV	Veh	260.9	3.2×	142.5	2.1×
	Rap	217.7	2.6×	69.5	1.6×
	%TGI	16.6%		51.2%	
GOLPH3	Veh	536.6	5.8×	223.0	2.5×
	Rap	102.5	1.9×	38.0	1.3×
	%TGI	80.9%		81.9%	

Figure 5 | The *in vivo* GOLPH3 growth advantage is abrogated by treatment with rapamycin. a, b, Mice containing tumours of the melanoma cell lines WM239A (**a**) and 1205LU (**b**) transduced with empty vector (EV; left panels) or *GOLPH3* (right panels), were treated with vehicle or rapamycin ($6.0\ mg\ kg^{-1}$) at two-day increments after treatment onset (tumour baseline volume $\sim 100\ mm^3$). Growth curves are plotted as the mean change in tumour volume relative to the baseline starting volume for each group. Error bars indicate ± s.e.m. for biological replicates. %TGI indicates the percentage tumour growth inhibition at time course end-point. **c,** Table summarizing data from the rapamycin-treatment xenograft studies in **a** and **b** at the same time point (day 8, after four doses). Rap, rapamycin; Veh, vehicle. Note 1205LU-GOLPH3 tumours treated with vehicle grew 2.5× in size during the 8 days of treatment. In comparison, WM239A-GOLPH3 tumours grew 5.8× in size during the same period of 8 days. The %TGI in these two cohorts of tumours was similar, at 81.9% and 80.9%, respectively, indicating that the growth rate did not impact on the response to rapamycin.

Discussion

In this study, integrative analyses of genome-wide copy number and expression data, coupled with reinforcing knockdown and overexpression assays *in vitro* and *in vivo*, led to the identification of GOLPH3 as a bona fide oncoprotein frequently targeted for copy-number gain/amplification in diverse human cancers. A suspect role for the Golgi apparatus in regulating cancer-relevant signalling has been speculated on the basis of the observation that some cytoplasmic membrane oncoproteins, such as RAS, can functionally signal when temporally present at the Golgi apparatus[24]. However, proteins such as GOLPH3 that are predominantly localized to the TGN have not been directly linked on a genetic level to cancer; therefore, GOLPH3 represents a first-in-class Golgi oncoprotein. Mechanistically, we show that the enhanced activation of mTOR signalling represents a molecular basis for the oncogenic activity of GOLPH3. In this light, enhanced and sustained mTOR activation *in vivo* would be expected to confer a significant growth advantage to cancer cells, a probable basis for increased *GOLPH3* gene copy number or expression in a large fraction of human cancers.

Although we show through genetic and pharmacological studies that the activation of mTOR is essential for the tumorigenicity of GOLPH3, the precise structural and biochemical basis for its activity remains to be fully determined. Our molecular data on the physical interaction between GOLPH3 and the retromer complex, which is responsible for protein trafficking between endosomes and the TGN[16], for the first time, to our knowledge, genetically implicates this biological process in cancer. This is consistent with recent reports on the essential role of the retromer and retrograde transport in regulation of the wntless receptor and proper secretion of the Wnt morphogen[25], which is important in both normal and neoplastic development. Along the same line, depletion of Vps35 (also known as CG5625) in *Drosophila* inhibited endocytosis of RTKs with concomitant alterations in downstream signalling[26]. Taken together, these data raise the possibility that GOLPH3 might function with VPS35 and the retromer to regulate receptor recycling of key molecules, thereby influencing downstream signalling through mTOR.

It has recently been discovered that Vps74, the yeast homologue of GOLPH3, is required for proper docking and localization of glycosyltransferases to the Golgi apparatus[27,28]. Protein glycosylation is one of the most prevalent forms of post-translational modification, and altered glycosylation is a hallmark feature of cancers[29]. It is noteworthy that glycosylation is known to be important for growth-factor activation of transmembrane receptors, because glycosylation mediates receptor sorting, ligand binding and endocytosis[29,30]. Thus, it is plausible that human GOLPH3 might serve a similar function in glycosyltransferase docking as in *S. cerevisiae*, and therefore might influence the downstream mTOR signalling response through its effect on membrane RTKs.

The PI3K–AKT–mTOR signalling cascade is activated in nearly all cancers and hence represents an intense focus for cancer drug development. However, the clinical response to rapamycin and its analogues has been feeble[11]. Our demonstration of the role of GOLPH3 in activating mTOR signalling and conferring increased sensitivity to rapamycin in preclinical setting raises the possibility that GOLPH3 expression level or gene copy-number status may predict sensitivity to mTOR inhibitors. Indeed, end-point analysis of our preclinical treatment studies showed that rapamycin was significantly more effective against xenograft tumours expressing high levels of GOLPH3 ($P = 0.0268$; 1205LU-GOLPH3 versus 1205LU-empty-vector tumour volumes at end-point; Supplementary Fig. 5b), thereby suggesting that GOLPH3 levels may be a positive predictor of rapamycin sensitivity. Although the predictive value of GOLPH3 as a biomarker remains to be demonstrated, the formulation of this hypothesis highlights the importance of mechanistic insights beyond functional and clinicopathological validation in the translation of cancer genomics.

METHODS SUMMARY

Array-CGH profiles of melanoma, NSCLC and colon adenocarcinoma were analysed as previously described[31]. FISH analyses on TMAs of nine tumour types using 5p13 and CEP1 centromeric probes determined frequency of copy number gain. Real-time quantitative genomic PCR in informative samples containing the 5p13 CNA defined a MCR containing four genes, and integration with NSCLC tumour and cell line transcriptome profiles identified MCR-resident genes with copy-number-correlated expression patterns. Oncogenic activity was assayed *in vitro* (proliferation, MEF transformation, and anchorage-independent growth in soft agar) and *in vivo* (subcutaneous tumour growth) through loss-of-function (by RNAi targeting *GOLPH3*) and gain-of-function (by overexpression of GOLPH3) in primary (MEF, HMEL) and transformed mouse and human cells. The yeast two-hybrid screen was performed using a human cDNA brain library with a *GOLPH3* bait construct. Serum- and EGF-induced mTOR and AKT activity was measured with pS6K(Thr 389) and pAKT(Ser 473) as readout by western blotting. Total mTOR and pS6K(Thr 389) were used to quantify relative protein expression level on tumour tissue microarrays by AQUA to correlate expression with copy-number status determined by FISH. Rapamycin sensitivity was determined by drug treatment in *in vivo* xenograft assays.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 19 October 2008; accepted 7 May 2009.

1. Manning, B. D. & Cantley, L. C. AKT/PKB signaling: navigating downstream. *Cell* **129**, 1261–1274 (2007).
2. Wullschleger, S., Loewith, R. & Hall, M. N. TOR signaling in growth and metabolism. *Cell* **124**, 471–484 (2006).

3. Shima, H. *et al.* Disruption of the p70^{s6k}/p85^{s6k} gene reveals a small mouse phenotype and a new functional S6 kinase. *EMBO J.* **17,** 6649–6659 (1998).
4. Montagne, J. *et al. Drosophila* S6 kinase: a regulator of cell size. *Science* **285,** 2126–2129 (1999).
5. Oldham, S., Montagne, J., Radimerski, T., Thomas, G. & Hafen, E. Genetic and biochemical characterization of dTOR, the *Drosophila* homolog of the target of rapamycin. *Genes Dev.* **14,** 2689–2694 (2000).
6. Zhang, H., Stallock, J. P., Ng, J. C., Reinhard, C. & Neufeld, T. P. Regulation of cellular growth by the *Drosophila* target of rapamycin dTOR. *Genes Dev.* **14,** 2712–2724 (2000).
7. Fingar, D. C. & Blenis, J. Target of rapamycin (TOR): an integrator of nutrient and growth factor signals and coordinator of cell growth and cell cycle progression. *Oncogene* **23,** 3151–3171 (2004).
8. Guertin, D. A. & Sabatini, D. M. Defining the role of mTOR in cancer. *Cancer Cell* **12,** 9–22 (2007).
9. Yang, Q. & Guan, K. L. Expanding mTOR signaling. *Cell Res.* **17,** 666–681 (2007).
10. Abraham, R. T. & Wiederrecht, G. J. Immunopharmacology of rapamycin. *Annu. Rev. Immunol.* **14,** 483–510 (1996).
11. Sabatini, D. M. mTOR and cancer: insights into a complex relationship. *Nature Rev. Cancer* **6,** 729–734 (2006).
12. Garraway, L. A. *et al.* Integrative genomic analyses identify MITF as a lineage survival oncogene amplified in malignant melanoma. *Nature* **436,** 117–122 (2005).
13. Wu, C. C. *et al.* GMx33: a novel family of trans-Golgi proteins identified by proteomics. *Traffic* **1,** 963–975 (2000).
14. Bell, A. W. *et al.* Proteomics characterization of abundant Golgi membrane proteins. *J. Biol. Chem.* **276,** 5152–5165 (2001).
15. Snyder, C. M., Mardones, G. A., Ladinsky, M. S. & Howell, K. E. GMx33 associates with the *trans*-Golgi matrix in a dynamic manner and sorts within tubules exiting the Golgi. *Mol. Biol. Cell* **17,** 511–524 (2006).
16. Bonifacino, J. S. & Hurley, J. H. Retromer. *Curr. Opin. Cell Biol.* **20,** 427–436 (2008).
17. Xie, M. W. *et al.* Insights into TOR function and rapamycin response: chemical genomic profiling by using a high-density cell array method. *Proc. Natl Acad. Sci. USA* **102,** 7215–7220 (2005).
18. Camp, R. L., Chung, G. G. & Rimm, D. L. Automated subcellular localization and quantification of protein expression in tissue microarrays. *Nature Med.* **8,** 1323–1327 (2002).
19. Fingar, D. C., Salama, S., Tsou, C., Harlow, E. & Blenis, J. Mammalian cell size is controlled by mTOR and its downstream targets S6K1 and 4EBP1/eIF4E. *Genes Dev.* **16,** 1472–1487 (2002).
20. Burnett, P. E., Barrow, R. K., Cohen, N. A., Snyder, S. H. & Sabatini, D. M. RAFT1 phosphorylation of the translational regulators p70 S6 kinase and 4E-BP1. *Proc. Natl Acad. Sci. USA* **95,** 1432–1437 (1998).
21. Isotani, S. *et al.* Immunopurified mammalian target of rapamycin phosphorylates and activates p70 S6 kinase α *in vitro. J. Biol. Chem.* **274,** 34493–34498 (1999).
22. Hresko, R. C. & Mueckler, M. mTOR·RICTOR is the Ser^{473} kinase for Akt/protein kinase B in 3T3–L1 adipocytes. *J. Biol. Chem.* **280,** 40406–40416 (2005).
23. Sarbassov, D. D., Guertin, D. A., Ali, S. M. & Sabatini, D. M. Phosphorylation and regulation of Akt/PKB by the rictor-mTOR complex. *Science* **307,** 1098–1101 (2005).
24. Chiu, V. K. *et al.* Ras signalling on the endoplasmic reticulum and the Golgi. *Nature Cell Biol.* **4,** 343–350 (2002).
25. Eaton, S. Retromer retrieves wntless. *Dev. Cell* **14,** 4–6 (2008).
26. Korolchuk, V. I. *et al. Drosophila* Vps35 function is necessary for normal endocytic trafficking and actin cytoskeleton organisation. *J. Cell Sci.* **120,** 4367–4376 (2007).
27. Schmitz, K. R. *et al.* Golgi localization of glycosyltransferases requires a Vps74p oligomer. *Dev. Cell* **14,** 523–534 (2008).
28. Tu, L., Tai, W. C., Chen, L. & Banfield, D. K. Signal-mediated dynamic retention of glycosyltransferases in the Golgi. *Science* **321,** 404–407 (2008).
29. Ohtsubo, K. & Marth, J. D. Glycosylation in cellular mechanisms of health and disease. *Cell* **126,** 855–867 (2006).
30. Takahashi, M., Tsuda, T., Ikeda, Y., Honke, K. & Taniguchi, N. Role of N-glycans in growth factor signaling. *Glycoconj. J.* **20,** 207–212 (2004).
31. Maser, R. S. *et al.* Chromosomally unstable mouse tumours have genomic alterations similar to diverse human cancers. *Nature* **447,** 966–971 (2007).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank R. DePinho for critical reading of the manuscript, and L. Cantley, as well as members of the Chin laboratory, for helpful discussion. We thank H. Ying for assistance with confocal microscopy. The rabbit polyclonal antibody against human GOLPH3 was kindly provided by J. J. Bergeron of McGill University. Mouse monoclonal antibody against human GOLPH3, C19, was generated at the Dana-Farber/Harvard Cancer Center Monoclonal Antibody Core Facility. K.L.S. is at present supported by a Postdoctoral Fellowship from the American Cancer Society (PF-07-039-01-CSM), and K.L.S. and O.K. were previously supported by a National Institutes of Health (NIH) Training Grant appointment in the Department of Dermatology at Brigham and Women's Hospital (5-T32-AR07098-31). K.-K.W. was supported by a Program of Research Excellence (SPORE) grant (P50 CA090578) and NIH grants (R01 AG2400401; R01 CA122794). The AQUA immunofluorescence study was supported by a grant from the NIH to D.L.R. (RO-1 CA 114277). This work is primarily supported by grants from the NIH to L.C. (RO1 CA93947; P50 CA93683).

Author Contribution O.K. identified *GOLPH3* as an oncogene target of 5p13; K.L.S. performed oncogene validation and mechanistic studies; E.I. and A.P. performed TMA FISH analysis; H.R.W. and D.E.F. performed the HMEL anchorage-independent growth assay; S.D. and J.W. assisted with immunofluorescence assays and provided technical support; M.W. performed the co-immunoprecipitation assays; V.A. and D.L.R. performed AQUA analysis; S.C. assisted with cell culture studies; A.S. aided phospholipid and biochemical analyses; M.-C.L. and K.-K.W. performed rapamycin xenograft assays; Y.X. performed computational analyses; T.F. and J.H. provided technical support; L.C. supervised the experiments and data interpretation; K.L.S. and L.C. wrote the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addresses to L.C. (lynda_chin@dfci.harvard.edu).

METHODS

Cell lines. All cell lines were propagated at 37 °C and 5% CO_2 in humidified atmosphere in RPMI 1640 medium (Invitrogen) supplemented with 10% heat-inactivated FBS. CRL-5889, SK-MEL-5, 1205LU, A549 and 293T cells were obtained from the American Type Culture Collection. Sbcl2, WM239A and hTERT/CDK4(R24C)/p53DD/BRAF(V600E) melanocytes (HMEL) have been described before[12,32].

Plasmids, retroviral transduction, and siRNA transfection. The retroviral HA–GOLPH3 expression clone, pBABE-HA-GOLPH3, was constructed by subcloning PCR-generated *GOLPH3* cDNA (NCBI accession NM_022130) into pBABE-puro-HA (Addgene). The *GOLPH3* siRNA-resistant construct pBABE-HA-GOLPH3 (siRES), which encodes wild-type GOLPH3 protein with nucleotide sequence mutated to resist siRNA against *GOLPH3* (siRNA 3), was constructed by site-directed mutagenesis of the pBABE-HA-GOLPH3 vector using primers 5′-TGTATGTTAATTGAATTAGCATTGAGGGGTAGATTGC AACTAGAGGCTTGTGGAATGAGACG-3′, and 5′-CGTCTCATTCCACAA GCCTCTAGTTGCAATCTACCCCTCAATGCTAATTCAATTAACATACA-3′. pEF-Dest51-GOLPH3, pLenti4/TO/V5-DEST-GOLPH3 and pLenti6/V5/DEST-VPS35 were constructed via gateway recombination cloning (Invitrogen) into pEF-Dest51, pLenti4/TO/V5-DEST and pLenti6/V5/DEST (Invitrogen), respectively, using pDONR223-GOLPH3 and pDONR223-VPS35 entry clones (Center for Cancer Systems Biology at the Dana-Farber Cancer Institute) according to the manufacturer's protocol. The yeast bait construct, pGBKT7-GOLPH3, was constructed by inserting the *GOLPH3* fragment from pBABE-HA-GOLPH3 into the pGBKT7 bait vector (Clontech). The tet-inducible GOLPH3 cell line, HMEL-tet-GOLPH3, was created using the T-Rex lentiviral expression system (Invitrogen) and pLenti4/TO/V5-DEST-GOLPH3 according to the manufacturer's protocol. Lentivirus and retrovirus were prepared by co-transfecting 293T cells with the above-mentioned vector backbones and standard virus packaging systems for subsequent collection of viral supernatants. All overexpression studies were performed with newly transduced stable cells lines. In HMEL-tet-GOLPH3 cells, GOLPH3 expression was stimulated by the addition of 2 $\mu g\,ml^{-1}$ doxycycline for 48 h for all assays.

For siRNA experiments, cells were seeded to approach 80% confluence at the time of transfection with Lipofectamine2000 (Invitrogen) according to the manufacturer's protocol. Transfections were performed with Block-It (Invitrogen) or siRNA (Dharmacon) targeting *GOLPH3* (siRNA pool, M-006414-00; siRNA 1, D-006414-01; siRNA 2, D-006414-02; siRNA 3, D-006414-03; siRNA 4, D-006414-04), *SUB1* (siRNA pool, M-009815-00) or non-targeting control (NT siRNA: D-001210-02-20). Cells were incubated 48–72 h before collection for all assays. Unless otherwise indicated, GOLPH3 knockdown experiments were conducted using siRNA 3. For soft agar colony formation and proliferation assays, cells were plated 24 h after siRNA transfection.

Cross-tumour array-CGH and expression analysis. Cross-tumour array-CGH analysis of malignant melanoma, NSCLC and colon adenocarcinoma was performed as described[31] using melanoma and colon adenocarcinoma data previously submitted to the Gene Expression Omnibus (accession numbers GSE7606 and GSE7604, respectively). The NSCLC data set has been previously described[33] and can be found at http://genomic.dfci.harvard.edu/. The numbers of 5p13 amplifications identified in array-CGH profiles were: melanoma, 6 (2 focal) present among 88 tumour profiles; NSCLC, 18 (1 focal) present among 67 profiles (15 cell lines and 52 tumours); colorectal cancer, 4 present among 81 profiles (38 cell lines and 43 tumours). For expression analysis of MCR-resident genes, NSCLC expression data accompanying array-CGH profiles was analysed. Among the 42 samples with both array-CGH 22K profiles and Affymetrix HGU133plus2 profiles available, 14 contained the 5p13 amplification event. Expression values for each gene (values from several probes for the same gene were averaged) from two groups (with or without amplification) were compared by two-sample *t*-test, and the significance level was adjusted with Bonferroni correction.

TMA-FISH. The following tissue microarrays were purchased from Cybrdi: CC04-01-004 lung carcinoma, CC11-11-005 ovarian carcinoma, CC08-01-002 breast carcinoma, CC05-21-001 colon adenocarcinoma, CC03-01-003 liver carcinoma and CC19-11-007 prostate carcinoma. The multiple myeloma tissue microarray was from TriStar. PA802 pancreatic carcinoma and ME1001 melanoma tissue microarrays were from US Biomax. FISH was prepared following standard protocols. BAC RP11-437P15 was used to mark the region of gain at 32 Mb on chromosome 5 (5p13). The centromere-specific CEP1 probe (Abbott Laboratories) served as a ploidy reference. FISH signal evaluation and acquisition were performed manually using filter sets and software developed by Applied Spectral Imaging. A signal to reference ratio greater than 1.5 was considered as gain; a ratio above $\geq$2.5 was considered as a high amplification level.

TMA-IHC and AQUA. Arrays were deparaffinized with xylene, rehydrated and antigen-retrieved by pressure cooking for 20 min in citrate buffer (pH 6). Slides were pre-incubated with 0.3% BSA in 0.1 M Tris-buffered saline (TBS, pH 8) for 30 min at room temperature. Lung cancer TMAs were then incubated overnight with a cocktail of either the mTOR primary antibody (1:200; rabbit monoclonal, clone 7C10, Cell Signaling Technology) and a mouse monoclonal anti-human cytokeratin antibody (1:100 in BSA/TBS; clone AE1/AE3, M3515, Dako) or the pS6K(Thr 389) primary antibody (1:200; mouse monoclonal, clone 1A5, Cell Signaling Technology) and a wide-spectrum rabbit anti-cow cytokeratin antibody (1:100 in BSA/TBS; Z0622, Dako). For the melanoma cohort, a mouse monoclonal S100 antibody (15E2E2, BoGenex) and a rabbit polyclonal S100 antibody (Z0311, Dako) (both 1:100 in BSA/TBS-Tween) were used instead of the mouse and rabbit cytokeratin, respectively. This was followed by a 1-h incubation with Alexa 546-conjugated goat anti-mouse secondary antibody (A11003, Molecular Probes) diluted 1:100 in rabbit EnVision reagent (Dako), and Alexa 546-conjugated goat anti-rabbit secondary antibody (A11010, Molecular Probes) diluted 1:100 in mouse EnVision reagent (Dako), for mTOR and pS6K, respectively. Cyanine 5 (Cy5) directly conjugated to tyramide (Perkin-Elmer) at a 1:50 dilution was used as the fluorescent chromagen for target detection. Prolong mounting medium (ProLong Gold, Molecular Probes) containing DAPI was used to identify tissue nuclei. Serial sections of a smaller TMA consisting of 30 lung cancer specimens (NSCLC 'test' array) were stained aside both cohorts to confirm assay reproducibility. H1299 and A431 cells were used as positive controls as indicated by the manufacturer. Negative control sections, in which the primary antibody was omitted, were used for each immunostaining run.

AQUA allows exact measurement of protein concentration within subcellular compartments, as described in detail elsewhere[18]. In brief, a series of high-resolution monochromatic images were captured by the PM-2000 microscope. For each histospot, in- and out-of-focus images were obtained using the signal from the DAPI, cytokeratin and S100-Alexa 546 (for lung cancer and melanoma, respectively) and mTOR/pS6K-Cy5 channels. mTOR and pS6K were measured using a channel with emission maxima above 620 nm, to minimize tissue autofluorescence. Tumour was distinguished from stromal and non-stromal elements by creating a tumour 'mask' from the cytokeratin and S100 signal for lung cancer and melanoma specimens, respectively. This created a binary mask (each pixel being either 'on' or 'off') on the basis of an intensity threshold set by visual inspection of histospots. The AQUA score of the protein of interest in each subcellular compartment was calculated by dividing the signal intensity (scored on a scale from 0 to 255) by the area of the specific compartment. Specimens with less that 5% tumour area per spot were not included in automated quantitative analysis for not being representative of the corresponding tumour specimen.

For statistical analysis, Pearson's correlation coefficient (r) was used to assess the correlation between log-normalized mTOR and pS6K AQUA scores, as well as the same scores on serial cuts of the NSCLC test array. Evaluation of the interarray reproducibility did not show significant differences between serial sections of the NSCLC test array (Pearson's $r = 0.95$, $P < 0.0001$). Ratios of cytoplasmic-to-nuclear expression for mTOR and pS6K were used to normalize for individual variation between groups, and were compared to *GOLPH3* gene copy number. The association between mTOR, pS6K AQUA scores and *GOLPH3* gene copy number (as determined by 5p13 FISH) with clinicopathologic parameters (age, gender, histological type and tumour differentiation) were analysed using Spearman's rank test. *GOLPH3* copy-number status was also binarized on the basis of FISH signals as normal ($<$1.5) and gain ($\geq$1.5), and correlated as non-continuous variable with all other parameters using Pearson's correlation.

Quantitative PCR and copy number. Genomic DNA was prepared from cell lines and melanomas with the Wizard kit (Promega), and total RNA was isolated using Trizol reagent (Invitrogen) and RNeasy columns (Qiagen). DNA copy numbers and relative expression levels were determined by real-time PCR using SYBR green I (Qiagen) detection chemistry, and the Stratagene MX3000p detection system according to the manufacturer's protocol. The comparative cycle threshold method was used to quantify target gene or mRNA copy numbers in the samples. For quantification of gene copy numbers, we compared copy numbers in tumour DNA to copies in normal human control DNA (Promega). The DNA copy number normalization reference was Line-1 DNA copy number.

Anchorage-independent growth. Soft-agar assays were performed on 6-well plates in triplicate. For each well, 1×10^4 cells were mixed thoroughly in cell growth medium containing 0.4% SeaKem LE agarose (Fisher) in RPMI plus 10% FBS, followed by plating onto bottom agarose prepared with 0.65% agarose in RPMI and 10% FBS. Each well was allowed to solidify and subsequently covered in 1 ml RPMI and 10% FBS, which was refreshed every 4 days. Colonies were stained with 0.05% (w/v) iodonitrotetrazolium chloride (Sigma) and scanned at

1,200 dots per inch (d.p.i.) using a flatbed scanner, followed by counting and two-tailed *t*-test calculation using Prism 4 (Graphpad).

Proliferation assays. Proliferation assays were performed on 12-well plates in triplicate using 7×10^3 (A549) or 1×10^4 (CRL-5889 and 1205LU) cells per well. Cells were fixed in 10% formalin in PBS, and stained with crystal violet at 24 h increments starting after cell adherence (T_0). At the conclusion of the assay, crystal violet was extracted using 10% acetic acid and measured at assessed for absorbance at 595 nm (A_{595}).

Focus formation assay. *Ink4a/Arf*-deficient primary MEFs were plated in DMEM containing 10% FBS at a density of 8×10^5 cells per 10 cm, 16 h before transfection. For RAS cooperation, 1.5 μg HRAS(Val 12) vector was co-transfected with 6.5 μg pEF-Dest51-LacZ control vector, MYC or pEF-Dest51-GOLPH3 using Lipofectamine2000 (Invitrogen) following the manufacture's instructions. The total amount of transfected DNA was kept constant at 7.5 μg, and transfections were done in duplicate three times. At 48 h after transfection, each transfected 10-cm plate was equally split into three 10-cm plates and incubated for 10 days, during which media was refreshed twice. Cells were washed, fixed in 10% formalin and stained with Giemsa solution (Sigma) for 10 min at room temperature for foci quantification. Two-tailed *t*-test calculations were performed using Prism 4 (Graphpad).

Yeast two-hybrid interaction screening. A pre-transformed human fetal brain cDNA library (Clontech) was screened (1×10^6 clones) using the AH109 yeast reporter strain and the Matchmaker Two-Hybrid System 3 (Clontech), according to the manufacturer's instructions. Plasmid DNA from 118 potential positive clones was isolated after transformation into *Escherichia coli* strain DH5α, followed by DNA sequencing using the provided prey vector-specific primers. Informative sequencing data was obtained for 102 of the 118 clones, 44 of which contained partial to full-length coding sequence, and were further considered for downstream analysis. GOLPH3-dependency was assessed by serial loss passaging the positive clones in the absence of bait vector selection, followed by replica plating VPS35-expressing AH109 cells to SC-LT (to confirm absence of the GOLPH3 bait vector) and to SC-L-H-A+XαGAL (to confirm loss of reporter activation).

Xenograft studies. WM239A and A549 cells transduced with either empty vector or HA–GOLPH3 retrovirus (pBABE-HA-GOLPH3) were stably selected and subcutaneously implanted in female nude animals (Taconic) at 1.0×10^6 and 2.5×10^6 cells per site on both flanks, respectively, mixed 1:1 with Matrigel (BD Bioscience). WM239A- and A549-derived tumours were isolated at 22 and 45 days, respectively, and measured for tumour volume. Two-tailed *t*-test calculations were performed using Prism 4 (Graphpad).

For *in vivo* rapamycin studies, melanoma 1205LU and WM239A cells verified to stably-express empty vector or GOLPH3 were subcutaneously implanted into female nude animals (Taconic) at 5.0×10^5 cells per site on both flanks. Tumours were allowed to reach approximately 100 mm^3, after which time animals were randomized into separate cohorts for treatment with vehicle (5% (v/v) PEG400, 5% (v/v) Tween-80) or rapamycin (6 mg kg^{-1}; LC Laboratories; 50 mg ml^{-1} stock prepared in 100% ethanol and diluted fresh in vehicle solvent for treatment), by intraperitoneal injection every other day. Tumour volumes and body weights were measured after drug administration. Tumour volume was determined by measuring in two directions with vernier calipers and formulated as tumour volume = (length × width^2)/2. Growth curves were plotted as the mean change in tumour volume for each group, in which the mean change indicates change in tumour volume at a given time point minus the tumour volume at time of initial dose administration. End-point scatter plots were plotted as a proportion of final dose end-point tumour volume over respective tumour baseline starting volume. Percentage tumour growth inhibition was determined as $(1 - (T/N)) \times 100$, in which T is the mean change in tumour volume of the treated group and N is the mean change in tumour volume of the control group at the assay end-point. Two-tailed *t*-test calculations were performed using Prism 4 (Graphpad).

Cell-size determinations. A Becton Dickinson FACScan flow cytometer with Cell Quest software was used to determine relative changes in cell size as assessed by forward-scatter differences. A549 cells were seeded into 60-mm dishes for siRNA transfection the next day as indicated above, followed by incubation overnight. Cells were split to 10-cm dishes, with or without 25 nM rapamycin (EMD Bioscience), and collected for flow cytometry by ethanol fixation 60 h after transfection. For FACS analysis, 10,000 single cells were collected and assessed for forward scatter.

Immunoblotting, immunofluorescence and co-immunoprecipitation analyses. For EGF stimulation assays, cells were serum starved for 24 h, followed by treatment as indicated with 100 ng ml^{-1} EGF (Invitrogen). For immunoblotting, cells were washed twice in PBS and lysed using RIPA buffer (Boston BioProducts) containing 1 mM PMSF, 1× Protease Inhibitor Cocktail (Sigma) and 1× Phosphatase inhibitor (Calbiochem) for separation on NuPAGE 4–12% Bis-Tris gels (Invitrogen), and blotted onto PVDF (Millipore). The following antibodies were used for immunoblotting following the manufacturer's recommendations: pS6K(Thr 389) (1:1,000; Cell Signaling), S6K (1:1,000; Cell Signaling), pAKT(Ser 473) (1:1,000; Cell Signaling), AKT (1:1,000; Cell Signaling), mTOR(Ser 2481) (1:1,000; Cell Signaling), alpha-tubulin (1:20,000; Sigma), vinculin (1:5,000; Santa Cruz), p4E-BP1(Thr 37/46), (1:1,000; Cell Signaling), pPDK1(Ser 241) (1:1,000; Cell Signaling), PTEN (1:750; Cell Signaling), phosphorylated p27Kip(Thr 157) (also known as CDKN1B) (1:500; R&D Systems), pMEK1/2(Ser 217/221) (1:1,000; Cell Signaling), phosphorylated-p44/42(Thr 202/Tyr 204) (1:1,000; Cell Signaling), V5 (1:5,000; Invitrogen) and HA (1:1,000; Cell Signaling). Rabbit anti-GOLPH3 (JJB; 1:1,000) used in Supplementary Fig. 2b was a gift from J. J. Bergeron, McGill University. All other GOLPH3 immunoblotting was performed using mouse anti-GOLPH3 (C19; 1:1,000) prepared by the Dana-Farber/Harvard Cancer Center Monoclonal Antibody Core Facility.

For co-immunoprecipitation studies, parental A549 cells or 293T cells cotransfected with the combinations of pBABE-HA-GOLPH3, pLenti6/V5/DEST-VPS35 or the corresponding empty vectors were used. Cytosolic lysates from these A549 cells or transfectants were prepared with hypotonic buffer (10 mM Tris, pH 7.6, 10 mM KCl, 5 mM MgCl2, 1% NP40 and protease inhibitors) 48 h after transfection (for 293T cells), and protein lysates were adjusted to 50 mM Tris, pH 7.6, and 150 mM NaCl for immunoprecipitation using either anti-V5, anti-HA or anti-GOLPH3 antibody overnight at 4 °C with rocking. Protein G beads (Roche) were added to the lysate–antibody mix following the manufacture's recommendations, and incubated for a further 3 h at 4 °C with rocking. Immunoprecipitates were washed three times for 20 min with either low stringent buffer for anti-V5 immunoprecipitation (PBS plus 0.1% NP40 and 0.05% nadeoxycholate), or high stringent buffer for anti-HA and anti-GOLPH3 immunoprecipitation (RIPA buffer with 500 mM NaCl). Immunoprecipitation complexes were eluted by the addition of SDS loading buffer after centrifugation, and resolved on NuPAGE 4–12% Bis-Tris gels (Invitrogen) for immunoblotting analysis for GOLPH3 (anti-HA or anti-GOLPH3) and VPS35 (anti-V5) .

For immunofluorescence studies, cells were cultured on coverslips, followed by fixation for 15 min at room temperature in 4% paraformaldehyde in PBS, permeabilization for 10 min at room temperature in 0.1% Triton X-100 in PBS and blocking for 1 h at room temperature in 10% goat or donkey serum in PBS. Slides were incubated for 1 h at room temperature with the following antibodies: mouse anti-HA (1:100; Cell Signaling), rabbit anti-TGN46 (1:500; Abcam), goat anti-VPS35 (1:200; Abcam) and mouse anti-GOLPH3 (1:300). Slides were stained for 1 h at room temperature with the corresponding Alexa Fluor secondary antibodies (Invitrogen). Coverslips were stained with DAPI (Sigma) and mounted with mounting medium. Microscopic images were obtained using a Nikon inverted Ti microscope equipped with Yokogawa spinning disk confocal/TIRF system and a Hamamatsu Orca ER firewire digital CCD camera, using constant exposure times for each channel in individual experiments. Images were compiled and false-coloured with Adobe Photoshop using identical settings for each colour. Magnification was ×100 unless otherwise indicated.

32. Satyamoorthy, K. *et al.* Melanoma cell lines from different stages of progression and their biological and molecular analyses. *Melanoma Res.* **7** (suppl. 2), S35–S42 (1997).
33. Tonon, G. *et al.* High-resolution genomic profiles of human lung cancer. *Proc. Natl Acad. Sci. USA* **102,** 9625–9630 (2005).

ARTICLES

Reconstitution of Rab- and SNARE-dependent membrane fusion by synthetic endosomes

Takeshi Ohya[1], Marta Miaczynska[1]†, Ünal Coskun[1], Barbara Lommer[1], Anja Runge[1], David Drechsel[1], Yannis Kalaidzidis[1,2] & Marino Zerial[1]

Rab GTPases and SNAREs (soluble *N*-ethylmaleimide-sensitive factor attachment protein receptors) are evolutionarily conserved essential components of the eukaryotic intracellular transport system. Although pairing of cognate SNAREs is sufficient to fuse membranes *in vitro*, a complete reconstitution of the Rab–SNARE machinery has never been achieved. Here we report the reconstitution of the early endosomal canine Rab5 GTPase, its key regulators and effectors together with SNAREs into proteoliposomes using a set of 17 recombinant human proteins. These vesicles behave like minimal 'synthetic' endosomes, fusing with purified early endosomes or with each other *in vitro*. Membrane fusion measured by content-mixing and morphological assays requires the cooperativity between Rab5 effectors and cognate SNAREs which, together, form a more efficient 'core machinery' than SNAREs alone. In reconstituting a fusion mechanism dependent on both a Rab GTPase and SNAREs, our work shows that the two machineries act coordinately to increase the specificity and efficiency of the membrane tethering and fusion process.

An outstanding question in the field of molecular cell biology is the functional relationship between Rab GTPases and SNAREs, key components of the intracellular trafficking apparatus[1–7]. Rab GTPases regulate the membrane recruitment and activity of tethering factors bringing together membranes compatible for fusion[2,3,8,9]. Membrane fusion is thought to be mediated by the formation of energetically stable *trans*-SNARE complexes engaging vesicle v-SNAREs (mostly R-SNAREs) and target t-SNAREs (mostly Q-SNAREs), enabling the closely apposed membranes to overcome the free energy barrier and fuse[4,5,10–12]. Because SNAREs can fuse membranes when reconstituted in proteoliposomes[4,5,10], they were proposed to be the core components of membrane fusion[10,13] and compartment specificity[5]. The original fusion reactions were inefficient, and required long times and non-physiological concentrations of SNAREs[5,10,13]. However, when SNAREs are in an assembled state, the transition from membrane docking to lipid mixing occurs in milliseconds[14,15]. Truncation of SNAREs enhances the efficiency and speed of membrane fusion[16,17], indicating that SNAREs contain regulatory domains that require other factors enabling their fusogenic activity. Moreover, other studies questioned the role of SNAREs as sole determinants of membrane fusion specificity[18]. This argues in favour of additional components regulating membrane fusion and compartmental specificity. Rab GTPases and their effectors are primary candidates for such a role.

Consistent with a function in determining the structural and functional identity of organelles, Rab GTPases and their effectors have a much narrower compartmental distribution[2,19,20] than SNAREs. On endosomes, assembly of the Rab5 domain involves a complex cascade of molecular interactions involving regulators of the nucleotide cycle (for example, rabaptin-5–rabex-5, a Rab5 effector–guanine nucleotide exchange factor (also termed rabaptin–RAB GEF 1) complex[21]) and Rab effectors. An important Rab5 effector is the phosphatidylinositol-3-OH kinase (PI(3)K) hVPS34 (also known as PIK3C3)–PIK3R4 (also known as p150) (refs 22, 23), which regulates the synthesis of phosphatidylinositol-3-phosphate (PtdIns(3)P)—a hallmark of early endosomes. Both Rab5 and PtdIns(3)P serve as binding sites for effectors required for membrane tethering and fusion (rabaptin-5–rabex-5 (ref. 21), early endosome antigen 1 (EEA1; ref. 8), rabenosyn-5 (also known as ZFYVE20)–hVPS45 (ref. 24), rabankyrin-5 (also known as ANKFY1; ref. 25)) through functional interactions with SNAREs[26] and intracellular motility[27]. Rab5-domain assembly and function therefore depends on cooperativity among Rab5 effectors and between these and SNAREs[2,28]. Here we reconstituted Rab5-domain assembly *in vitro*, with the intent of recapitulating membrane tethering and fusion, more faithfully and efficiently than SNAREs alone.

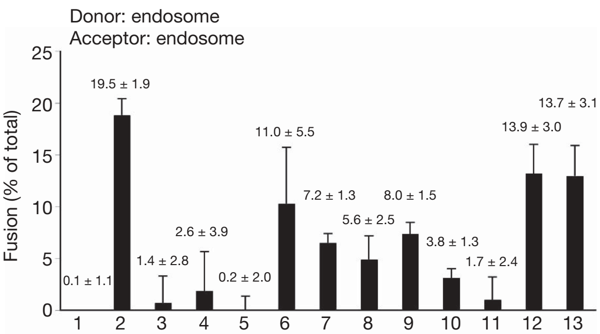
Rab effectors in early endosome fusion *in vitro*

A set of recombinant proteins, including Rab5, its interacting proteins and effectors[2] as well as the t- or Q-SNAREs syntaxin 13 (also known as syntaxin 12), VTI1A and syntaxin 6 and the v- or R-SNARE VAMP4 (refs 18, 26, 29) responsible for early endosomes fusion, were expressed either in *Escherichia coli* or Sf+ cells using the baculovirus expression system and purified (some partially, Supplementary Fig. 1 and Supplementary Table 1). First, we ensured that the purified proteins were active in an *in vitro* homotypic endosome fusion measured using a content mixing assay[21]. In brief, 'donor' endosomes purified from HeLa cells after internalization of biotinylated transferrin were mixed with 'acceptor' endosomes containing endocytosed anti-transferrin antibodies (Supplementary Fig. 2a). The biotinylated immunocomplexes were captured on streptavidin-coated plates and revealed using ruthenium labels that emit light when electrochemically stimulated (Meso Scale Discovery, MSD; see Methods). This electro-chemiluminescence system has the advantages of high sensitivity and dynamic range, as well as minimal background signals (Supplementary Fig. 2b, c). Addition of unlabelled transferrin in the

[1]Max Planck Institute of Molecular Cell Biology and Genetics, Pfotenhauerstrasse 108, 01309, Dresden, Germany. [2]A. N. Belozersky Institute of Physico-Chemical Biology, Moscow State University, Moscow 119899, Russia. †Present address: International Institute of Molecular and Cell Biology (IIMCB), Ks. Trojdena 4, 02-109 Warsaw, Poland.

reaction prevents the formation of labelled immunocomplexes resulting from leakage of markers upon membrane lysis, as observed for proteoliposomes[30,31] or yeast vacuoles[32]. Solubilization of membranes without unlabelled transferrin yields the total amount of immunocomplexes achievable in the assay (100% signal). The background in homotypic endosome fusion was estimated by solubilizing membranes without incubation and in the presence of unlabelled transferrin to be $4.6 \pm 1.0\%$ of the total signal and subtracted from each measurement. As a result, some values of the fusion activity presented in the graphs are negative within the range of 0–2%, the inherent variability of the system.

No significant fusion signal was detected when donor and acceptor endosomes were incubated with the ATP-regenerating system alone. The addition of cytosol ($3\ \mathrm{mg\ ml^{-1}}$) from HeLa cells enhanced fusion, which reached ~19% of the total signal (Fig. 1, compare lanes 1 and 2). We next attempted to substitute cytosol with the purified recombinant proteins added at concentrations similar to those provided by cytosol (see Supplementary Table 1). The concomitant addition of 50 nM recombinant prenylated Rab5–Rab guanine dissociation inhibitor (RabGDI) complex[33], 30 nM rabaptin-5–rabex-5 (ref. 21), 5 nM hVPS34–PIK3R4 (ref. 22), 100 nM EEA1 (ref. 8), rabenosyn-5–hVPS45 (ref. 24), rabankyrin-5 (ref. 25), 500 nM *N*-ethylmaleimide-sensitive factor (NSF) and soluble NSF attachment protein alpha (α-SNAP, also known as NAPA) required for the priming of *cis*-SNARE complexes[34] stimulated endosome fusion to an extent similar to cytosol (~19% versus 14%; Fig. 1, lanes 2 and 13). Selectively omitting individual components gave indications regarding their requirement for the fusion reaction (Fig. 1, lanes 3–12). For example, NSF and α-SNAP were essential. In contrast, omission of the Rab5–RabGDI complex produced relatively high fusion activity (Fig. 1, lane 6), although clearly lower than that stimulated by cytosol, probably due to residual native Rab5 and effectors on the purified endosomes (Supplementary Fig. 3), consistent with previous studies[22,35]. Exclusion of rabaptin-5–rabex-5, hVPS34–PIK3R4, EEA1 or rabenosyn-5–hVPS45 also generated minimal (~2–8%) fusion activity. Only removal of rabankyrin-5 did not decrease the fusion. Nevertheless, the recombinant protein is functional because



	1	2	3	4	5	6	7	8	9	10	11	12	13
Cytosol		+											
NSF				+		+	+	+	+	+	+	+	+
α-SNAP			+			+	+	+	+	+	+	+	+
Rab5–GDI			+	+	+		+	+	+	+	+	+	+
Rabaptin-5–rabex-5			+	+	+	+		+	+	+	+	+	+
hVPS34–PIK3R4			+	+	+	+	+		+	+	+	+	+
EEA1			+	+	+	+	+	+		+		+	+
Rabenosyn-5–hVPS45			+	+	+	+	+	+	+			+	+
Rabankyrin-5			+	+	+	+	+	+	+	+	+		+

Figure 1 | Rab5 effectors and SNARE priming factors can substitute cytosol in the homotypic fusion between early endosomes *in vitro*. Homotypic early endosome fusion was carried out in the absence (lane 1) or presence of cytosol (lane 2) or the indicated combination of recombinant proteins at the concentration specified in Supplementary Table 1 (lanes 3–13) at 37 °C for 25 min. Fusion efficiency is expressed as percentage of the total possible fusion signal (see Methods). The precise values are indicated above each column. All data show the average of three independent experiments and s.e.m.

100 nM rabankyrin-5 stimulated the fusion reaction with cytosol by 25% (not shown, see ref. 25). We conclude that, with the exception of rabankyrin-5, all recombinant factors of the Rab5 machinery tested here are necessary and, together, sufficient for substituting cytosol in the early endosome fusion *in vitro*.

Recruitment of Rab5 and effectors on proteoliposomes

Next, we attempted to replace purified endosomes with proteoliposomes harbouring SNAREs complemented by the Rab5 machinery. To deliver prenylated Rab5 by RabGDI onto the membranes, we also implanted recombinant His_6-tagged prenylated Rab acceptor 1 (PRA1, also known as RABAC1), because this protein acts as a GDI displacement factor (GDF) primarily on Rab9, but also on Rab5 (ref. 36). Proteoliposomes were prepared by mixing recombinant proteins solubilized in detergent with lipids extracted from endosome fractions (to closely mimic the endosomal lipid composition) followed by dialysis of detergent and flotation on Histodenz density gradient (see Methods). Importantly, this protocol yielded a SNARE-to-phospholipid molar ratio similar to that of purified native endosomes (see syntaxin 13 in Fig. 2a, compare lanes 1–5 with 6). We then monitored the recruitment of Rab5, EEA1 and rabenosyn-5–hVPS45 onto these proteoliposomes by western blot analysis. Figure 2a shows that RabGDI delivered Rab5 on proteoliposomes containing PRA1 (compare lane 1 with 4). However, delivery of Rab5 alone was not sufficient for efficient membrane recruitment of EEA1 and rabenosyn-5, consistent with the dependence on Rab5 activation by GDP/GTP exchange (lane 2). Note that the amount of Rab5 recruited on the proteoliposomes in this case was low, presumably due to the counteracting extraction by RabGDI. Addition of the rabaptin-5–rabex-5 complex increased the membrane recruitment of Rab5 but not of its effectors (lane 3). Further addition of hVPS34–PIK3R4 enhanced the recruitment of EEA1 and rabenosyn-5 on the proteoliposomes (lane 4), to a similar extent (EEA1), if not better (rabenosyn-5), than the recruitment on purified endosomes (lane 6). The dependency on hVPS34–PIK3R4 could be bypassed by incorporating exogenous PtdIns(3)P (0.01 mol %, lane 5) into the proteoliposomes. These results corroborate previous proposals from studies *in vitro* and *in vivo* arguing that membrane assembly of the Rab5 machinery is based on effector cooperativity[2,28].

Whereas purified early endosomes efficiently recruit Rab5 effectors from cytosol *in vitro*, proteoliposomes containing PRA1 but without SNAREs were stripped of the recruited Rab5 effectors when incubated with cytosol (not shown). Presumably, cytosol provides endogenous Rab5 effectors but also factors counter-acting the assembly process, for example, negatively acting upon Rab5 and PtdIns(3)P such as RabGDI, Rab5 GTPase-activating proteins, PtdIns(3)P-phosphatases and other molecules competing for Rab5 binding[8]. Given that both EEA1 and rabenosyn-5–hVPS45 complex can bind syntaxin 13 and syntaxin 6 on endosomes[24,26,37,38], we tested whether such interactions may contribute to the membrane recruitment of Rab5 effectors. PRA1 proteoliposomes containing either cognate syntaxin 13, VTI1A and syntaxin 6 or non-cognate t-SNAREs (where syntaxin 13 was replaced by syntaxin 7) were first incubated with the full set of recombinant Rab5 effectors except rabankyrin-5, and then further incubated with and without cytosol. Figure 2b shows that proteoliposomes containing cognate t-SNAREs recruited similar amounts of Rab5, EEA1 and rabenosyn-5–hVPS45 both in the presence and absence of cytosol (compare lane 1 with 2). In contrast, non-cognate syntaxin 7, VTI1A and syntaxin 6 proteoliposomes poorly recruited Rab5 and its effectors (especially rabenosyn-5) and incubation with cytosol completely solubilized them (lanes 3–4). Moreover, proteoliposomes containing only PRA1 and syntaxin 13 alone displayed substantial sensitivity to cytosol (lanes 5–6). These data indicate that the correct cognate SNARE complexes contribute to the recruitment and stabilization of Rab5 effectors on the membrane.

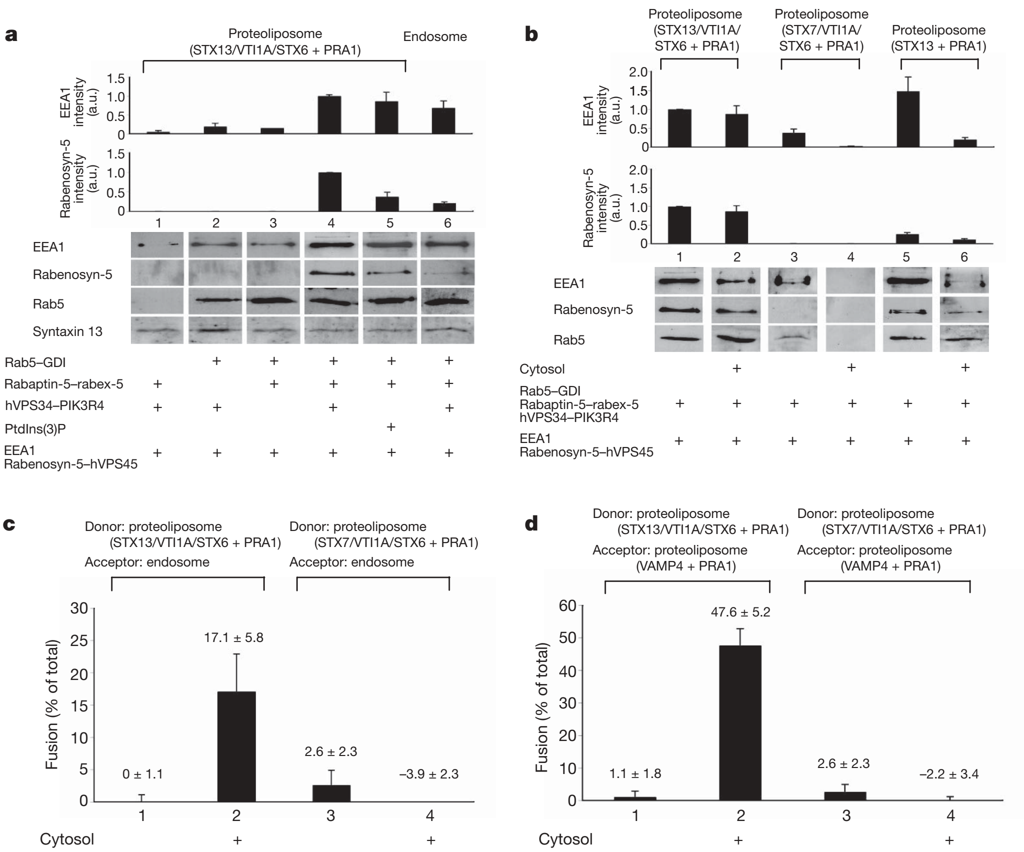


Figure 2 | Recruitment of Rab5, EEA1 and rabenosyn-5 on proteoliposomes. **a**, Proteoliposomes with a similar phospholipid and syntaxin 13 (STX13) content as early endosomes (as control, lane 6) were incubated with the indicated proteins (see Methods). Membrane-associated proteins were detected by western blotting, quantified and normalized to 1 (arbitrary units, a.u.) with the values in lane 4. **b**, Cognate t-SNAREs are necessary for the stable membrane recruitment of EEA1 and rabenosyn-5. Proteoliposomes with different SNARE sets were incubated with recombinant proteins with or without further incubation with cytosol. Membrane association was measured as in **a**. **c**, **d**, Cognate t-SNAREs are necessary for cytosol-dependent proteoliposome fusion. Membrane fusion of either cognate (lanes 1–2) or non-cognate (lanes 3–4) t-SNARE proteoliposomes with endosomes (**c**) or v-SNARE proteoliposomes (**d**) with or without cytosol was measured as in Fig. 1. Histograms show the average of three independent experiments and s.e.m.

Previous studies used proteoliposomes with a SNARE: phospholipid ratio ranging between 1:22–1:200 (refs 10, 13) and 1:500–1:2,000 (ref. 39). In our experiments, the proteoliposomes contained similar amounts of SNAREs as early endosomes (see Methods and Supplementary Fig. 4). We estimated that ~2.5 picomoles syntaxin 13 were present in the early-endosome-enriched fraction containing ~5.5 nanomoles phospholipids, resulting in a ~1:2,000 ratio. A similar ratio (~1:10,000–1:2,000, depending on the preparation) was estimated for the proteoliposomes because the concentration of syntaxin 13 and phospholipids incorporated into these membranes matched that of endosomes. The copy number of syntaxin 13 would be ~4–20 per 100 nm liposome (see Methods), the same order of magnitude of SNAREs in synaptic vesicles[40]. The concentration of syntaxin 6 was also comparable to that of endosomal fractions whereas VTI1A was present at a somewhat lower level (see Supplementary Fig. 4). VAMP4 in endosomes and proteoliposomes was barely detectable by western blotting but was not in excess over the level in endosomes.

Fusion of proteoliposomes requires cytosolic factors

We then tested the fusion of the PRA1 and t-SNARE proteoliposomes with endosomes using the same assay measuring homotypic early endosome fusion (Fig. 1). In Fig. 2c, proteoliposomes harbouring syntaxin 13, VTI1A, syntaxin 6, PRA-1 and containing biotinylated transferrin in the lumen (referred to here as donor) fused with acceptor early endosomes (containing anti-transferrin antibodies) to a comparable degree (~17.1%, lane 2) as homotypic endosome fusion (Fig. 1, lane 2), and this fusion was cytosol-dependent (compare lane 1 with 2). In contrast, donor proteoliposomes harbouring syntaxin 7, VTI1A and syntaxin 6 yielded low fusion signals that were further reduced by addition of cytosol (lanes 3 and 4), indicating that membrane fusion requires the correct cognate early endosomal t-SNARE complex[18,29].

We next attempted to substitute the acceptor endosomes with proteoliposomes containing PRA1, the cognate v-SNARE VAMP4 and anti-transferrin antibodies in the lumen. As for the endosome fusion assay, the background in the proteoliposome assay, estimated to be 5.5 ± 1.1% of the total signal, was subtracted from each measurement. Strikingly, donor proteoliposomes harbouring PRA1, syntaxin 13, VTI1A and syntaxin 6 fused with very high efficiency with acceptor proteoliposomes containing PRA1 and VAMP4 but only in the presence of cytosol (48% versus 1.1%, Fig. 2d, lanes 1 and 2) and, again, this fusion activity was dependent on the correct t-SNARE complex (compare lane 2 with 4). Thus, in this content-mixing assay, cognate SNAREs alone do not efficiently support membrane fusion.

Proteoliposome fusion requires the Rab5 machinery and SNAREs

The dependence on cytosolic factors may be accounted for, at least partially, by the Rab5 machinery. To test this hypothesis directly, we attempted to reconstitute a fully synthetic fusion reaction, using only purified recombinant components (Fig. 3a). Donor proteoliposomes with PRA1, syntaxin 13, VTI1A and syntaxin 6 incubated with acceptor proteoliposomes (PRA1, VAMP4) again yielded a very low fusion signal (1.1%). Similarly, low signals (0.9–4.6%) were obtained with proteoliposomes containing cognate SNAREs but lacking PRA1 or containing PRA1 but lacking SNAREs. Strikingly,

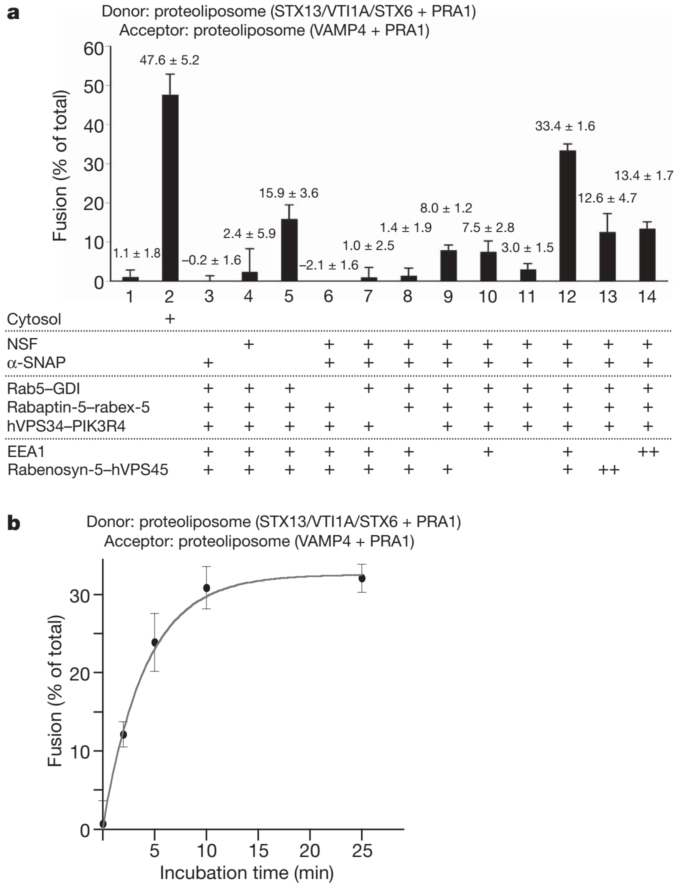


Figure 3 | The Rab5 machinery and SNAREs cooperatively promote membrane fusion. **a**, Donor proteoliposomes containing syntaxin 13, VTI1A, syntaxin 6 and PRA1 and acceptor proteoliposomes with VAMP4 and PRA1 were incubated without cytosol (lane 1), with cytosol (lane 2) or with the indicated recombinant proteins (lanes 3–14) at the concentration in Supplementary Table 1, as described in Methods. In lanes 13 and 14, '++' indicates the presence of 200 nM of proteins. Fusion was measured as in Fig. 1. **b**, Time course of fully synthetic proteoliposome membrane fusion. The fusion assay was carried out as in **a**, lane 12, for the indicated periods of time. Black dots, experimental data; red curve, exponential fit ($t_{1/2}$ = 2.8 ± 0.7 min). Graphs show the average of three independent experiments ± s.e.m.

the same proteoliposomes incubated with Rab5–RabGDI and the Rab5 effectors rabaptin-5–rabex-5 complex, hVPS34–PIK3R4, EEA1, rabenosyn-5–hVPS45 complex, and NSF and α-SNAP (at the concentration specified in Supplementary Table 1) at 37 °C for 25 min fused very efficiently (~33%) as shown in Fig. 3a, lane 12. The SNARE priming factors NSF and α-SNAP were required for maximal fusion signal (lanes 3–5). This is because the amount of content in the donor and acceptor liposomes allows for an increased fusion signal after multiple rounds of fusion. However, the fusion signal obtained in the absence of NSF and α-SNAP (lane 5) indicates that the t- and v-SNAREs on the separate proteoliposomes allow at least one cycle of membrane fusion without the need for priming[41]. Interestingly, omission of either α-SNAP or NSF was more inhibitory than removal of both (compare lanes 3 and 4 with 5). In the absence of NSF, surplus α-SNAP is known to bind to SNAREs non-productively[42]. NSF incorporates into high-molecular-weight oligomers with syntaxin 13, EEA1, rabaptin-5 and rabex-5 (ref. 26), but evidently without α-SNAP these complexes do not support membrane fusion. The reaction was dependent on Rab5 (lane 6), catalytic nucleotide exchange on Rab5 (lane 7) and GTP (not shown). Similar to the recruitment of Rab5 effectors (Fig. 2a), it also required the synthesis of PtdIns(3)P by hVPS34–PIK3R4 (lane 8). Note that the same complement of recombinant proteins as in lanes 6–8 supported relatively high endosome fusion (Fig. 1, lanes 6–8), probably due to residual effectors on the cellular membranes. Although either EEA1 or the rabenosyn-5–hVPS45 complex alone was necessary and sufficient for fusion (lanes 9–11), maximal fusion activity required the synergistic contribution of both Rab5 effectors (lane 12). Doubling the concentration of either effector to compensate for the loss of the other did not yield the fusion activity obtained when both were present (Fig. 3, lanes 13–14).

The fusion reaction was saturable and occurred with a $t_{1/2}$ of 2.8 ± 0.7 min (Fig. 3b). The kinetics were thus even faster than those reported for the homotypic fusion between early endosomes[35]. Altogether, these results indicate that we succeeded in reconstituting an efficient and specific proteoliposome-dependent fusion system recapitulating both the synergistic activity between components of the Rab5 machinery and the Rab- and SNARE-dependency.

Although all components in the reconstituted system are required, some exert a more regulatory role on others and, thus, can be bypassed. For example, similar to the recruitment of Rab5 effectors (Fig. 2a, b), incorporation of PtdIns(3)P into proteoliposomes bypassed the requirement for PI(3)K (Fig. 4a, lanes 3 and 4). Addition of 1 mM GTPγS in part bypassed the requirement for the rabaptin-5–rabex-5 (data not shown).

To validate the artificial proteoliposome system, we conducted a series of control experiments. For these, a 'standard full-set fusion assay' was defined as the reaction with donor and acceptor proteoliposomes and recombinant proteins as in Fig. 3a, lane 12.

First, we further tested the SNARE dependency by adding soluble syntaxin 13 or syntaxin 7 fragments lacking the transmembrane domain[26] to standard full-set fusion assays. Soluble syntaxin 13 inhibited fusion (21% of the control; Fig. 4b, lanes 1–2), whereas syntaxin 7 had a modest inhibitory effect (80% of the control, lanes 1 and 3). Altogether, these data indicate that the fusion between donor and acceptor membranes with cognate early endosome SNAREs specifically requires syntaxin 13.

Second, we verified that the immunocomplexes detected are not due to increased leakage of the luminal markers from tethered or aggregated proteoliposomes. Standard full-set fusion assays without unlabelled transferrin as the quencher were carried out, and the proteoliposomes separated into three fractions by Histodenz flotation (Fig. 4c, see Methods). As controls, donor and acceptor proteoliposomes were either directly separated without incubation or lysed after incubation before flotation. As expected, when proteoliposomes were detergent-solubilized most of the signal was detected in the bottom fraction. In contrast, most (~70%) of the cargo was detected in the top fraction for all donor, acceptor and proteoliposome products of the fusion reaction. These results indicate that the incubation does not increase content leakage from donor or acceptor proteoliposomes. Note that the luminal markers leaking from membranes are quenched by addition of excess unlabelled transferrin down to 1.1% (Fig. 3a, lane 1). In this way, we can exclude that the immunocomplexes detected in the standard full-set fusion assay are primarily due to leakage of cargo rather than membrane fusion. In fact, the signals detected are probably an under-estimate of the real fusion signal.

To rule out that fusion activity may be caused by impurities in the lipid extracts and to reconstitute membranes using a defined system, we generated artificial proteoliposomes consisting of synthetic and purified lipids (see Methods), mimicking the endosomal phospholipid content[43]. Such defined proteoliposomes with cognate t-SNAREs and PRA1 were able to fuse with equivalent lipid-defined proteoliposomes with v-SNARE and PRA1 in the presence of either cytosol or recombinant Rab5 machinery, NSF and α-SNAP (Fig. 4d).

Morphological assessment of proteoliposome fusion

Finally, to assess the fusion of proteoliposomes by an independent method, we measured the diameter and area of proteoliposomes by

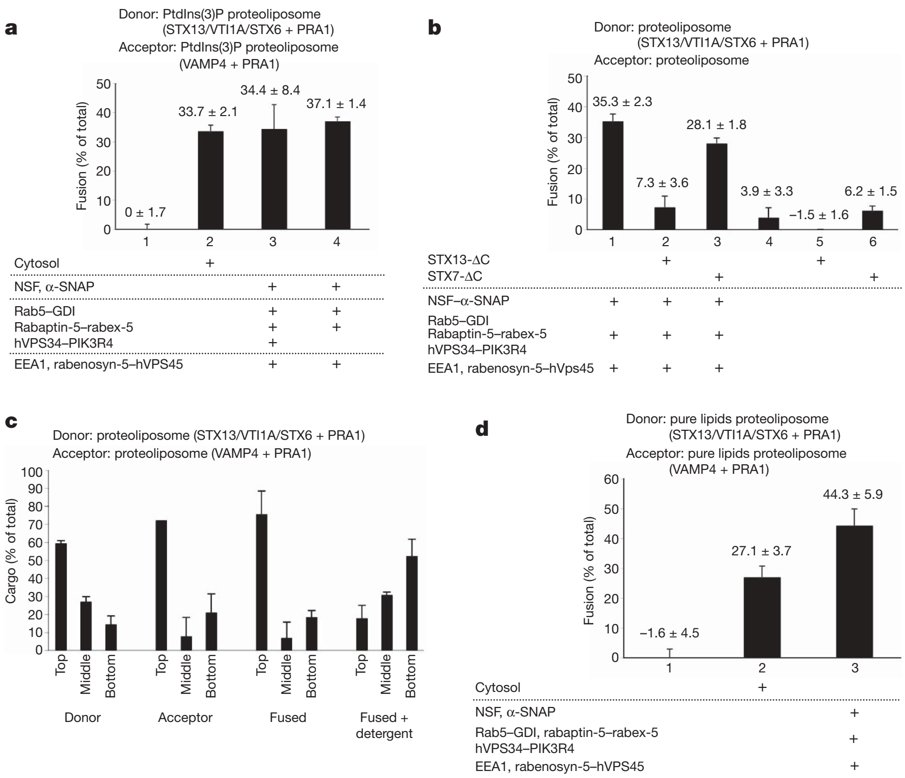


Figure 4 | Molecular requirements for membrane fusion. a, Exogenous PtdIns(3)P can bypass the requirement for PI(3)K. Proteoliposomes were supplemented with 0.01 mol % exogenous PtdIns(3)P and incubated in the absence or presence of cytosol or the indicated proteins. Fusion was measured as in Fig. 1. **b**, Requirement for SNAREs in membrane fusion assessed through the inhibitory effect of soluble syntaxin mutants. The fusion reaction (as in Fig. 3a, lane 12) was carried out in the absence or presence of 4 μM Stx13-ΔC (lanes 2 and 5) or Stx7-ΔC (lanes 3 and 6). **c**, Assessment of cargo leakage from proteoliposomes. Individual donor, acceptor or proteoliposomes incubated as in Fig. 3a, lane 12 (fused), without unlabelled transferrin were separated by Histodenz density gradient flotation and the cargo in each fraction was quantified. As a control, fused proteoliposomes were lysed before flotation (fused + detergent). **d**, Fusion between proteoliposomes reconstituted with synthetic lipids. Fusion assays with proteoliposomes consisting of pure lipids were performed as in Fig. 3a, lane 12. Data are average of three independent experiments and s.e.m.

negative staining and electron microscopy. A high number of proteins and/or protein concentrations could potentially produce membrane fragments or ruptured liposomes. We could observe, however, that donor proteoliposomes appeared as spherical vesicles (Fig. 5a). Their size ranged between ~30 and 500 nm, with most having an average diameter of ~50–100 nm (Fig. 5d). Acceptor proteoliposomes showed a similar profile (data not shown). Incubation of donor and acceptor proteoliposomes alone moderately increased the fraction of vesicles ranging from 250 to 500 nm (Fig. 5b, d), but not over 500 nm. However, incubation in the standard full-set fusion assay (Fig. 3a, lane 12) shifted the distribution profile of vesicles, yielding spherical proteoliposomes ranging between 0.5 and 1.5 μm in diameter (Fig. 5c, d). Further incubation times (for example, 24 h) did not yield significantly larger proteoliposomes over 2 μm (not shown). Using a simple mathematical model (see Supplementary Methods), we estimated that in the standard full-set fusion assay proteoliposomes undergo a minimum average of ~4.2 rounds of fusion compared with 0.17 rounds for PRA1 and SNARE-only proteoliposomes, consistent with the requirement of the Rab5 machinery, α-SNAP and NSF for maximal fusion (Fig. 3a).

We also verified that the enlarged proteoliposomes generated in the standard full-set fusion assay were indeed the products of bilayer fusion and not fusion intermediates, for example, hemifusion, by ammonium molybdenum staining and electron microscopy analysis. For donor, acceptor and fused proteoliposomes, we confirmed the bilayer membrane structure in proteoliposomes of a wide size range (Fig. 5e).

Discussion

In this study, we reconstituted the cooperative activity between the Rab5 and SNARE endosomal machinery in membrane tethering and fusion, using a combination of artificial vesicles and a set of 17 recombinant proteins. This is the first successful reconstitution of Rab-dependent membrane fusion showing the intimate interplay between Rab GTPases and SNAREs. Our results provide important mechanistic insights into membrane tethering and fusion. First, assembly of a functional Rab5 machinery requires the combinatorial activity of four different effectors (rabaptin-5–rabex-5, hVPS34–PIK3R4 PI(3)K, EEA1 and rabenosyn-5–hVPS45). Second, we could experimentally demonstrate the long-standing prediction that Rab proteins impart an additional layer of specificity on membrane tethering and fusion[2,44] to the pairing between cognate SNAREs[10]. Third, the fusion activity reconstituted here, as revealed by a bona fide content-mixing assay, used physiological concentrations of SNAREs, was even more efficient than that of early endosomes and required short incubation times. In our system, the early endosomal SNAREs proved to be very inefficient in fusing membranes (Fig. 3a), yielding signals close to background, arguing that Rab effectors together with SNAREs form a more efficient 'core machinery' than SNAREs alone.

Multiple activities may underlie the functional cooperativity between the Rab machinery and SNAREs in membrane fusion. First, one established function of Rab5 effectors is membrane tethering, as first demonstrated for EEA1 (ref. 8). This activity may help to stabilize cognate *trans*-SNAREs[26], increasing the probability to assemble fusion-competent complexes. In this respect, members of the yeast Sec1/eukaryote Munc-18 family are essential co-factors for SNAREs in membrane fusion[45,46], which for early endosomes is contributed by hVPS45 bound to rabenosyn-5 (ref. 24). Second, SNAREs contribute to the recruitment and stability of Rab5 effectors on the membrane. Third, Rab5 effectors may directly participate in membrane fusion.

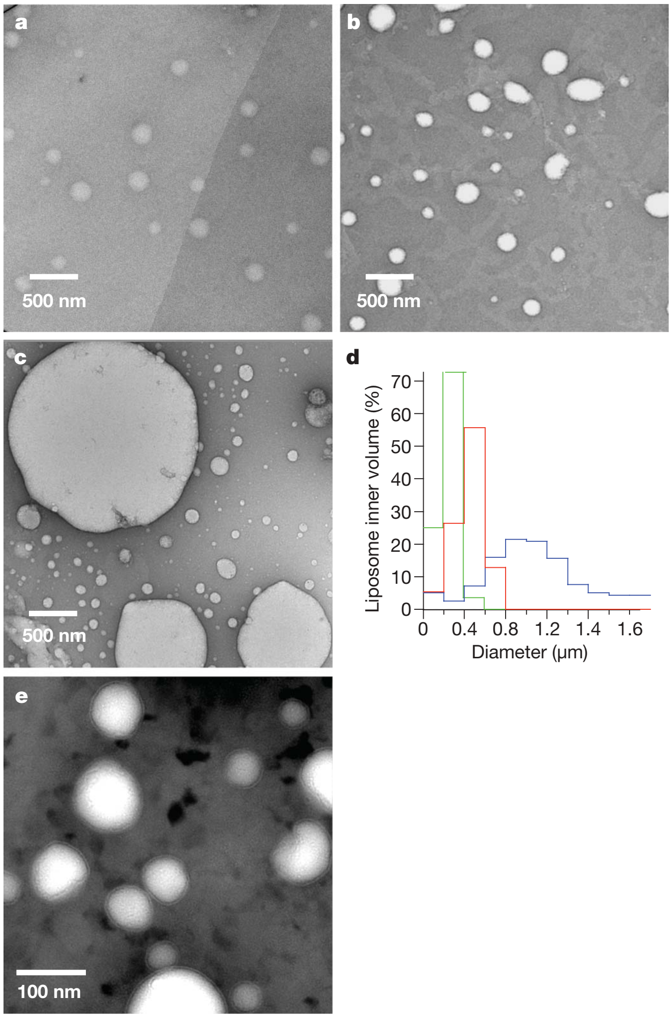


Figure 5 | Electron microscopic analysis of proteoliposomes stained with uranyl acetate (a–d) or ammonium molybdenum (e). a, Donor proteoliposomes. **b**, **c**, Donor and acceptor proteoliposomes incubated in the absence (**b**) or presence (**c**) of recombinant factors as in Fig. 3a, lane 12. **d**, Distribution of proteoliposome volumes in **a**–**c**. The histogram shows the distribution of proteoliposome volumes (as spheres calculated from the apparent diameters) normalized to 100% of the total volume of proteoliposomes measured under each experimental condition, from at least three independent preparations. Red, donor proteoliposomes as in **a**, $n = 628$; green, donor and acceptor proteoliposomes as in **b**, $n = 602$; blue, standard full-set fusion mixture of donor and acceptor proteoliposomes as in **c**, $n = 3{,}787$. **e**, Proteoliposomes after standard fusion reaction were stained with 4% ammonium molybdenum.

They may have a mechanical role by increasing lateral tension and inducing membrane curvature, as shown for synaptotagmin I (ref. 47), thus lowering the energy barrier necessary to convert the hemifusion stalk into an expanding fusion pore[12].

The reconstituted system developed here still lacks important components, such as members of the synaptotagmin family[47] and class C VPS–CORVET complex[48,49]. Some of the components used here *in vitro* may be replaced or complemented by others *in vivo*, contributing to the efficiency, specificity and regulation of membrane fusion. This synthetic biology approach could therefore be further developed towards the reconstitution of artificial organelles of increasing complexity and functional properties.

METHODS SUMMARY

Antibodies and recombinant proteins. Antibodies and the expression and purification of recombinant proteins in this Article are described in Supplementary Methods.

Proteoliposome preparation and purification. Extracted lipids from the endosome fraction of baby hamster kidney BHK cells or synthetic lipids were mixed with either PRA1 or SNARE proteins or both, in buffer containing biotinylated transferrin for t-SNARE proteoliposomes and sheep anti-transferrin antibodies for v-SNARE proteoliposomes. The mixtures were subjected to two steps of dialysis to form the proteoliposomes. Dialysed proteoliposomes were purified on a Histodenz gradient and used for protein binding and fusion studies. Details are provided in Methods.

Membrane recruitment assay. The purified proteoliposomes were incubated with recombinant proteins and the recruitment of Rab5 effectors was determined by quantitative western blotting.

Early endosome and proteoliposome fusion assays. The standard early endosome fusion assay was carried out as described previously[21]. See Methods for details. For proteoliposome fusion assays, t-SNAREs and v-SNARE proteoliposomes were purified by a two-step density-separation method, Histodenz flotation followed by sucrose gradient separation. Standard fusion assays with proteoliposomes were carried out as homotypic early endosome fusion assays except using donor proteoliposomes with acceptor proteoliposomes or endosomes (see Methods).

Electron microscopy visualization of proteoliposomes and quantification. Proteoliposomes were prepared on freshly glow-discharged carbon-coated copper grids stained with 1% uranyl acetate or 4% ammonium molybdenum. Images of the negatively stained proteoliposome specimen were collected on a TECNAI12 (FEI) electron microscope operating at 100 kV.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 13 February; accepted 5 May 2009.
Published online 20 May 2009.

1. Pfeffer, S. R. Transport-vesicle targeting: tethers before SNAREs. *Nature Cell Biol.* **1**, E17–E22 (1999).
2. Zerial, M. & McBride, H. Rab proteins as membrane organizers. *Nature Rev. Mol. Cell Biol.* **2**, 107–117 (2001).
3. Grosshans, B. L., Ortiz, D. & Novick, P. Rabs and their effectors: achieving specificity in membrane traffic. *Proc. Natl Acad. Sci. USA* **103**, 11821–11827 (2006).
4. Jahn, R. & Scheller, R. H. SNAREs—engines for membrane fusion. *Nature Rev. Mol. Cell Biol.* **7**, 631–643 (2006).
5. McNew, J. A. *et al.* Compartmental specificity of cellular membrane fusion encoded in SNARE proteins. *Nature* **407**, 153–159 (2000).
6. Rothman, J. E. & Sollner, T. H. Throttles and dampers: controlling the engine of membrane fusion. *Science* **276**, 1212–1213 (1997).
7. Cai, H., Reinisch, K. & Ferro-Novick, S. Coats, tethers, Rabs, and SNAREs work together to mediate the intracellular destination of a transport vesicle. *Dev. Cell* **12**, 671–682 (2007).
8. Christoforidis, S., McBride, H. M., Burgoyne, R. D. & Zerial, M. The Rab5 effector EEA1 is a core component of endosome docking. *Nature* **397**, 621–625 (1999).
9. Wang, L., Merz, A. J., Collins, K. M. & Wickner, W. Hierarchy of protein assembly at the vertex ring domain for yeast vacuole docking and fusion. *J. Cell Biol.* **160**, 365–374 (2003).
10. Weber, T. *et al.* SNAREpins: minimal machinery for membrane fusion. *Cell* **92**, 759–772 (1998).
11. Wickner, W. & Schekman, R. Membrane fusion. *Nature Struct. Mol. Biol.* **15**, 658–664 (2008).
12. Zimmerberg, J. & Gawrisch, K. The physical chemistry of biological membranes. *Nature Chem. Biol.* **2**, 564–567 (2006).
13. Schuette, C. G. *et al.* Determinants of liposome fusion mediated by synaptic SNARE proteins. *Proc. Natl Acad. Sci. USA* **101**, 2858–2863 (2004).
14. Fix, M. *et al.* Imaging single membrane fusion events mediated by SNARE proteins. *Proc. Natl Acad. Sci. USA* **101**, 7311–7316 (2004).
15. Liu, T. *et al.* SNARE-driven, 25-millisecond vesicle fusion *in vitro. Biophys. J.* **89**, 2458–2472 (2005).
16. Parlati, F. *et al.* Rapid and efficient fusion of phospholipid vesicles by the alpha-helical core of a SNARE complex in the absence of an N-terminal regulatory domain. *Proc. Natl Acad. Sci. USA* **96**, 12565–12570 (1999).
17. Pobbati, A. V., Stein, A. & Fasshauer, D. N- to C-terminal SNARE complex assembly promotes rapid membrane fusion. *Science* **313**, 673–676 (2006).
18. Brandhorst, D. *et al.* Homotypic fusion of early endosomes: SNAREs do not determine fusion specificity. *Proc. Natl Acad. Sci. USA* **103**, 2701–2706 (2006).
19. Sonnichsen, B. *et al.* Distinct membrane domains on endosomes in the recycling pathway visualized by multicolor imaging of Rab4, Rab5, and Rab11. *J. Cell Biol.* **149**, 901–914 (2000).
20. Barbero, P., Bittova, L. & Pfeffer, S. R. Visualization of Rab9-mediated vesicle transport from endosomes to the trans-Golgi in living cells. *J. Cell Biol.* **156**, 511–518 (2002).
21. Horiuchi, H. *et al.* A novel Rab5 GDP/GTP exchange factor complexed to Rabaptin-5 links nucleotide exchange to effector recruitment and function. *Cell* **90**, 1149–1159 (1997).

22. Christoforidis, S. *et al.* Phosphatidylinositol-3-OH kinases are Rab5 effectors. *Nature Cell Biol.* **1**, 249–252 (1999).
23. Shin, H. W. *et al.* An enzymatic cascade of Rab5 effectors regulates phosphoinositide turnover in the endocytic pathway. *J. Cell Biol.* **170**, 607–618 (2005).
24. Nielsen, E. *et al.* Rabenosyn-5, a novel Rab5 effector, is complexed with hVPS45 and recruited to endosomes through a FYVE finger domain. *J. Cell Biol.* **151**, 601–612 (2000).
25. Schnatwinkel, C. *et al.* The Rab5 effector Rabankyrin-5 regulates and coordinates different endocytic mechanisms. *PLoS Biol.* **2**, 1363–1380 (2004).
26. McBride, H. M. *et al.* Oligomeric complexes link Rab5 effectors with NSF and drive membrane fusion via interactions between EEA1 and syntaxin 13. *Cell* **98**, 377–386 (1999).
27. Hoepfner, S. *et al.* Modulation of receptor recycling and degradation by the endosomal kinesin KIF16B. *Cell* **121**, 437–450 (2005).
28. Del Conte-Zerial, P. *et al.* Membrane identity and GTPase cascades regulated by toggle and cut-out switches. *Mol. Syst. Biol.* **4**, 206 (2008).
29. Zwilling, D. *et al.* Early endosomal SNAREs form a structurally conserved SNARE complex and fuse liposomes with multiple topologies. *EMBO J.* **26**, 9–18 (2007).
30. Bhalla, A., Chicka, M. C., Tucker, W. C. & Chapman, E. R. Ca^{2+}-synaptotagmin directly regulates t-SNARE function during reconstituted membrane fusion. *Nature Struct. Mol. Biol.* **13**, 323–330 (2006).
31. Dennison, S. M., Bowen, M. E., Brunger, A. T. & Lentz, B. R. Neuronal SNAREs do not trigger fusion between synthetic membranes but do promote PEG-mediated membrane fusion. *Biophys. J.* **90**, 1661–1675 (2006).
32. Starai, V. J., Jun, Y. & Wickner, W. Excess vacuolar SNAREs drive lysis and Rab bypass fusion. *Proc. Natl Acad. Sci. USA* **104**, 13551–13558 (2007).
33. Ullrich, O., Horiuchi, H., Bucci, C. & Zerial, M. Membrane association of Rab5 mediated by GDP-dissociation inhibitor and accompanied by GDP/GTP exchange. *Nature* **368**, 157–160 (1994).
34. Mayer, A., Wickner, W. & Haas, A. Sec18p (NSF)-driven release of Sec17p (α-SNAP) can precede docking and fusion of yeast vacuoles. *Cell* **85**, 83–94 (1996).
35. Rybin, V. *et al.* GTPase activity of Rab5 acts as a timer for endocytic membrane fusion. *Nature* **383**, 266–269 (1996).
36. Sivars, U., Aivazian, D. & Pfeffer, S. R. Yip3 catalyses the dissociation of endosomal Rab–GDI complexes. *Nature* **425**, 856–859 (2003).
37. Simonsen, A., Gaullier, J. M., D'Arrigo, A. & Stenmark, H. The Rab5 effector EEA1 interacts directly with syntaxin-6. *J. Biol. Chem.* **274**, 28857–28860 (1999).
38. Ungermann, C., Price, A. & Wickner, W. A new role for a SNARE protein as a regulator of the Ypt7/Rab-dependent stage of docking. *Proc. Natl Acad. Sci. USA* **97**, 8889–8891 (2000).
39. Mima, J. *et al.* Reconstituted membrane fusion requires regulatory lipids, SNAREs and synergistic SNARE chaperones. *EMBO J.* **27**, 2031–2042 (2008).
40. Takamori, S. *et al.* Molecular anatomy of a trafficking organelle. *Cell* **127**, 831–846 (2006).
41. Weber, T. *et al.* SNAREpins are functionally resistant to disruption by NSF and alphaSNAP. *J. Cell Biol.* **149**, 1063–1072 (2000).
42. Collins, K. M. & Wickner, W. T. Trans-SNARE complex assembly and yeast vacuole membrane fusion. *Proc. Natl Acad. Sci. USA* **104**, 8755–8760 (2007).
43. Kobayashi, T. *et al.* A lipid associated with the antiphospholipid syndrome regulates endosome structure and function. *Nature* **392**, 193–197 (1998).
44. Pfeffer, S. & Aivazian, D. Targeting Rab GTPases to distinct membrane compartments. *Nature Rev. Mol. Cell Biol.* **5**, 886–896 (2004).
45. Dulubova, I. *et al.* Munc18–1 binds directly to the neuronal SNARE complex. *Proc. Natl Acad. Sci. USA* **104**, 2697–2702 (2007).
46. Shen, J. *et al.* Selective activation of cognate SNAREpins by Sec1/Munc18 proteins. *Cell* **128**, 183–195 (2007).
47. Martens, S., Kozlov, M. M. & McMahon, H. T. How synaptotagmin promotes membrane fusion. *Science* **316**, 1205–1208 (2007).
48. Rink, J., Ghigo, E., Kalaidzidis, Y. & Zerial, M. Rab conversion as a mechanism of progression from early to late endosomes. *Cell* **122**, 735–749 (2005).
49. Peplowska, K. *et al.* The CORVET tethering complex interacts with the yeast Rab5 homolog Vps21 and is involved in endo-lysosomal biogenesis. *Dev. Cell* **12**, 739–750 (2007).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We are grateful to K. Simons and B. Hoflack for discussions, to C. Stroupe and W. Wickner for sharing unpublished information, and to G. Marsne, I. Baines, W. Huttner, K. Simons, C. Stroupe and W. Wickner for critical reading of the manuscript. We acknowledge support by the systems biology network HepatoSys of the German Ministry for Education and Research (BMBF, grant 0313082J), the EU Integrated Project EndoTrack, the DFG and the Max Planck Society (including the Max Planck Partner Group grant to M.Z. and M.M.). T.O. was supported by The Nakatomi Foundation.

Author Contributions M.M. conducted the initial studies and tested the recombinant proteins in endosome fusion and the membrane recruitment of Rab5 and its effectors on proteoliposomes, and B.L. further developed such a proteoliposome system. D.D. and A.R. established several of the protocols of purification of recombinant proteins. T.O. completed the development of these procedures and conducted all biochemical experiments on membrane fusion reported in this study. Ü.C. performed the electron microscopy analysis, Y.K. did the statistical analysis and the mathematical model of membrane fusion, and M.Z. conceived and directed the project and wrote the manuscript with the help of T.O., M.M. and Y.K.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to M.Z. (zerial@mpi-cbg.de).

METHODS

Antibodies. Rabbit polyclonal anti-EEA1 and anti-rabenosyn-5 antibodies were previously described[24], as well as anti-Rab5 and anti-syntaxin 13 (ref. 26).

Expression and purification of recombinant proteins. The expression and purification of recombinant proteins in this Article are described in Supplementary Methods.

Proteoliposome preparation. Early-endosome-enriched fractions from baby hamster kidney (BHK-21) cells were prepared as described[43]. Lipids were extracted from using chloroform/methanol (2:1, v/v), dried by nitrogen stream and dissolved in liposome buffer, 20 mM Hepes/OH, pH 7.5, 1% CHAPS and 500 mM KCl for t-SNARE (syntaxin 6, VTI1A and syntaxin 13 or syntaxin 7) proteoliposomes, and the same buffer containing 300 mM KCl for v-SNARE (VAMP4) proteoliposomes. For the incorporation of exogenous PtdIns(3)P into proteoliposomes, 0.01 mol % of synthetic PtdIns(3)P was added to the lipid extract from the BHK endosome fraction in chloroform/methanol solution. The reconstitution of proteoliposomes with synthetic and purified phospholipids (Avanti Polar Lipids) followed the same protocol as above except that 96 nanomoles of PC (48 mol %), 50 nanomoles of PE (25 mol %), 18 nanomoles of SM (9 mol %), 18 nanomoles of PS (9 mol %) and 18 nanomoles of PI (9 mol %) were dissolved in 4 ml of chloroform/methanol (2:1, v/v). Either recombinant His_6-tagged PRA1 (final 500 nM) or SNARE proteins (each final 10 nM) or both were added to 120 μl of lipid solution (1 mM phospholipids) containing biotinylated transferrin (final 250 $\mu g\,ml^{-1}$) for t-SNARE proteoliposomes and sheep anti-transferrin antibodies (Scottish Antibody Production Unit, final 75 $\mu g\,ml^{-1}$) for v-SNARE proteoliposomes. Each mixture was incubated with gentle agitation in the dark at room temperature. Next, 120 μl of the t-SNARE mixture was first dialysed against 12 ml of dialysis-500 buffer (25 mM Hepes/OH, pH 7.5, 500 mM KCl) with 20 $\mu g\,ml^{-1}$ holo-transferrin (Sigma) for 12–15 h in the dark at 4 °C, followed by a second dialysis against 120 ml of dialysis-500 buffer for 4 h in the dark to remove CHAPS. In the case of v-SNARE proteoliposomes, the mixture was subjected to two steps of dialysis as above in dialysis-300 buffer (25 mM Hepes/OH, pH 7.5, 300 mM KCl), with the first step supplemented with 4.5 $\mu g\,ml^{-1}$ sheep anti-transferrin antibodies.

Proteoliposome purification and lipid quantification. Proteoliposomes were purified by Histodenz density gradient centrifugation. Dialysed t-SNARE proteoliposomes were mixed gently with 360 μl of 48% Histodenz buffer (48% Histodenz, 1 mM GTP, 100 μM $CaCl_2$ and 300 mM KCl in fusion buffer, see below) and then covered with 400 μl of 25% Histodenz buffer and 160 μl of Histodenz-free buffer, followed by centrifugation in a TLS-55 rotor (Beckman) at 100,000*g* for 1 h at 4 °C. Fractions of 40 μl were collected from the top to the bottom of the gradient and SNARE-enriched fractions, as determined by western blot analysis, at the interphase between the first and the second phases (fourth fraction from the top) were selected for protein-binding studies. The phospholipid concentration of proteoliposomes in this fraction was about ~2.5 mM, as quantified by the Bartlett assay[50].

Membrane recruitment assay. The protein recruitment assay was performed in the fusion buffer containing 12.5 mM Hepes/OH, pH 7.4, 1.5 mM magnesium oxaloacetate, 75 mM potassium acetate, 3 mM imidazole and 1 mM DTT[21]. The purified proteoliposomes (1 mM phospholipid) were incubated with recombinant proteins in the presence of an ATP-regeneration system, 1 μM $CaCl_2$ and 30 μM GTP in a total volume of 20 μl fusion buffer at 37 °C for 25 min. Recombinant proteins were adjusted to the final concentrations indicated in Supplementary Table 1. When the effect of cytosol was evaluated, samples were supplemented with HeLa cytosol (final concentration 3 $mg\,ml^{-1}$) and further incubated at 37 °C for 25 min. The reaction was arrested by transferring the sample on ice. Samples were then supplemented up to 35 μl with fusion buffer, followed by purification again by flotation on Histodenz density gradient consisting of 105 μl of 48%, 80 μl of 25% and 30 μl of Histodenz-free buffer, followed by fractionation as above. The floated fractions (top 50 μl) were used for testing the recruitment of Rab5 effectors as determined by quantitative western blotting using recombinant proteins as standards.

Membrane fusion assays. Standard early endosome fusion assay was carried out as previously described[21]. In brief, 'donor' and 'acceptor' endosomes were purified from HeLa cells that had internalized either biotinylated human transferrin or sheep anti-transferrin antibodies, respectively, for 5 min. Both endosome-enriched fractions were mixed with excess unlabelled transferrin (2 $mg\,ml^{-1}$), an ATP-regenerating system and incubated with 3 $mg\,ml^{-1}$ HeLa cytosol at 37 °C for 25 min (see Supplementary Fig. 2a). Membranes were solubilized in 2% Triton X-100 buffer at room temperature for 1 h and the resulting immunocomplexes between biotinylated transferrin and anti-transferrin antibodies were immobilized onto streptavidin-coated 96-well plates and detected by a rabbit anti-sheep secondary antibody coupled to MSD SULFO-TAG (ruthenium (II) tris-bipyridine) in an electrochemical reaction using the SECTOR Imager 6000 (MSD). In this system the electrochemical reaction is initiated when current is applied to the plates. In combination with the co-reactant tripropylamine, which is contained in the read buffer, the MSD SULFO-TAG emits light (620 nm) that is detected by a highly sensitive CCD camera. Background signals are minimal because the stimulus (electricity) is decoupled from the signal (light). Multiple excitation cycles of each label amplify the signal to enhance light levels and improve sensitivity. Detection limits may be as low as 10^6 molecules with a dynamic range of 6 logs. In our system, a few femtomoles of immunocomplexes could be detected (see Supplementary Fig. 2b, c). To quantify the maximal possible fusion signal (total), mixed acceptor and donor endosomes in each assay were solubilized in the absence of unlabelled transferrin and the resulting immunocomplex was quantified as described above. This value was considered as 100%. In this study, acceptor endosomes (final concentration in the assay, 80 μM phospholipids) and donor endosomes (final concentration, 40 μM phospholipids) were mixed in a total volume of 20 μl fusion buffer supplemented with 1 μM $CaCl_2$ and 30 μM GTP as a standard. When testing purified recombinant proteins instead of cytosol in the homotypic endosome fusion assay, the final concentration of each indicated recombinant protein was adjusted to the value indicated in Supplementary Table 1.

For the fusion assay with synthetic proteoliposomes, t-SNARE and v-SNARE proteoliposomes were purified by a two-step density-separation method. For t-SNARE proteoliposomes, the first step was done by flotation on a Histodenz gradient, in the same way as for the membrane recruitment assay. The top 240 μl fraction was sequentially mixed with an equal volume of 62% sucrose buffer (62% sucrose, 3 mM imidazole and 300 mM KCl). The proteoliposome mixture was covered with 200 μl of 35% sucrose buffer, 150 μl of 25% sucrose buffer and 50 μl of 8.5% sucrose buffer, followed by centrifugation in a TLS-55 rotor at 100,000*g* at 4 °C for 1 h. Fractions of 40 μl were collected from the top to the bottom of the gradient, and the SNARE-enriched fractions were determined as described above (usually seventh or eighth fraction from the top; Supplementary Fig. 4) and used as the 'donor' in the fusion assay. The purification of v-SNARE proteoliposomes was also carried out in the same way, except that KCl was omitted in both density gradient separation processes. The final fraction was used as the 'acceptor' in the fusion assay. The levels of SNARE proteins in the donor or acceptor proteoliposomes purified by sucrose density gradient were estimated by quantitative western blot analysis using recombinant proteins as standards and compared with those in endosomes (adjusted to obtain comparable amounts according to the phospholipid content). For example, we estimated that ~1.0 picomoles syntaxin 13 were present in the proteoliposomes containing ~2.2 nanomoles phospholipids, resulting in a molar ratio ranging between ~1:10,000 and 1:2,000, depending on the preparation and the error in the measurement. Assuming that a 40 nm vesicle contains ~7,000 phospholipids[40] covering 50% of the surface of the membrane (the other 50% being covered by transmembrane proteins and non-phospholipids, that is, cholesterol), we estimated that the copy number of syntaxin 13 ranges between ~4 and 20 in a 100 nm proteoliposome (correcting for 6.25 times the surface area).

Calibration experiments were conducted to find the optimal ratio between donor and acceptor proteoliposomes. First, the concentration of donor proteoliposomes (containing biotinylated transferrin) was adjusted to yield half maximal intensity of the total signal generated upon solubilization and incubation with excess of anti-transferrin antibodies. Similarly, the concentration of acceptor proteoliposomes (containing anti-transferrin antibodies) was adjusted to yield half maximal intensity of the total fusion signal obtained upon solubilization with donor proteoliposomes at the selected concentration. This corresponded to three times higher concentration of donors than acceptors. For this, donor proteoliposomes (final concentration in the assay, 75 μM phospholipids) and acceptor proteoliposomes (final concentration, 25 μM phospholipids) were mixed in the fusion assay. Standard fusion assays with proteoliposomes were also carried out with unlabelled transferrin, an ATP-regeneration system, 1 μM $CaCl_2$ and 30 μM GTP at 37 °C for 25 min. No addition or addition of 100 μM $CaCl_2$ did not result in significant differences in fusion signal (Supplementary Fig. 5).

For the experiments of proteoliposome–endosome fusion, donor proteoliposomes (final concentration, 75 μM phospholipids) were mixed with acceptor endosomes (final concentration, 80 μM phospholipids).

In the endosome fusion assay with cytosol, ~1,500 electrochemiluminescence (ECL) units of emission signal were detected as an endosome fusion activity and ~6,000 ECL units as a total using 0.8 nanomoles phospholipids in donor and 1.6 nanomoles phospholipids in acceptor endosomes. They corresponded to 4.2 femtomoles and 17.5 femtomoles of biotinylated transferrin and anti-transferrin antibodies, respectively (Supplementary Fig. 2). In the fusion between proteoliposomes, we obtained ~1,000 ECL units as a fusion activity in the presence of cytosol (Fig. 3a, lane 2) and ~2,000 ECL units as a total using 1.5 nanomoles phospholipids in donor and 0.5 nanomoles phospholipids in acceptor proteoliposomes. They corresponded to 1.5 femtomoles and 4.3

femtomoles of immunocomplex, respectively (Supplementary Fig. 2). Concerning the cargo, in endosomes, the estimated molecular ratio of biotinylated transferrin to phospholipids was ~1:90,000, and that of anti-transferrin antibody to phospholipids was ~1:45,000. The estimated molecular ratio of biotinylated transferrin to phospholipids in donor proteoliposomes was ~1:300,000, and that of anti-transferrin antibody in acceptor proteoliposomes was ~1:100,000. Soluble truncated mutants of syntaxin 13 (Stx13-ΔC: 1–234 amino acids) or syntaxin 7 (Stx7-ΔC: 1–217 amino acids) lacking the transmembrane domains were expressed and purified as previously described[26]. Soluble syntaxins (4 μM) were added to the standard fusion assay mixtures with endosome or proteoliposomes immediately before starting incubation at 37 °C for 25 min.

Proteoliposome flotation assay. The proteoliposome fusion reaction was carried out in the absence of unlabelled transferrin like a standard fusion assay with recombinant Rab5 machinery, NSF and α-SNAP proteins in the concentrations described in Supplementary Table 1. The reaction mixture (20 μl) was adjusted to 30 μl with fusion buffer and mixed with 90 μl of 48% Histodenz buffer, and then covered with 100 μl of 25% Histodenz buffer and 40 μl of Histodenz-free buffer, followed by centrifugation in a TLS-55 rotor (Beckman) at 100,000*g* at 4 °C for 1 h. The 'top' 120 μl, 'middle' 60 μl and the 80 μl at the 'bottom' were taken, solubilized in 2% Triton X-100 buffer and the resulting complex between biotinylated transferrin and anti-transferrin antibodies was quantified as described above. As control experiments, the same volume of donor and acceptor were adjusted to 30 μl with fusion buffer and subjected to Histodenz gradient flotation. Each donor fraction was incubated with 200 picomoles of anti-transferrin antibodies in 2% Triton X-100 at room temperature for 1 h. For acceptor membrane, 100 picomoles of biotinylated transferrin were pre-incubated with streptavidin immobilized in the MSD plate. Each acceptor fraction was also solubilized in 2% Triton X-100 buffer at room temperature for 15 min, and then further incubated with biotinylated transferrin pre-fixed on the MSD plate at room temperature for 1 h. The immunocomplex detection was carried out as described above. As a negative control, fused proteoliposomes were solubilized in 2% Triton X-100 immediately after the fusion reaction and subjected to Histodenz gradient flotation, followed by the quantification of their cargo in each fraction.

Electron microscopy visualization and quantification of proteoliposomes. Proteoliposomes were prepared on freshly glow-discharged carbon-coated copper grids stained with 1% uranyl acetate or 4% ammonium molybdenum. Images of the negatively stained proteoliposome specimen were collected on a TECNAI12 (FEI) electron microscope operating at 100 kV. Fusion reaction in the presence or absence of recombinant NSF, α-SNAP and Rab5 machinery proteins was carried out as a standard full-set fusion assay.

To provide an unbiased quantification, electron microscopy images were analysed using the MotionTracking software[48] to automatically recognize and statistically analyse the parameters of proteoliposomes. In brief, the images were smoothed by bilateral filter and proteoliposomes were identified by two algorithms: watershed algorithm for vesicles with diameter above 200 nm, and base-function fitting for objects smaller than 200 nm. The results of the identification were verified manually and are shown in Fig. 5d for 628 donor or acceptor proteoliposomes, donor and acceptor proteoliposomes containing PRA1 and cognate v- and t-SNAREs incubated at 37 °C for 25 min in the absence (602) and presence (3,787) of the full set of proteins.

Estimation of average number fusion rounds per proteoliposome. The number of fusion rounds was inferred from the mean volume of the proteoliposomes on the basis of the model described in the Supplementary Methods.

50. Bartlett, G. R. Colorimetric assay methods for free and phosphorylated glyceric acids. *J. Biol. Chem.* 234, 469–471 (1959).

LETTERS

Sodium salts in E-ring ice grains from an ocean below the surface of Enceladus

F. Postberg[1,2], S. Kempf[2,3], J. Schmidt[4], N. Brilliantov[5,6], A. Beinsen[7], B. Abel[7,8], U. Buck[9] & R. Srama[2]

Saturn's moon Enceladus emits plumes of water vapour and ice particles from fractures near its south pole[1–5], suggesting the possibility of a subsurface ocean[5–7]. These plume particles are the dominant source of Saturn's E ring[7,8]. A previous *in situ* analysis[9] of these particles concluded that the minor organic or siliceous components, identified in many ice grains, could be evidence for interaction between Enceladus' rocky core and liquid water[9,10]. It was not clear, however, whether the liquid is still present today or whether it has frozen. Here we report the identification of a population of E-ring grains that are rich in sodium salts (~0.5–2% by mass), which can arise only if the plumes originate from liquid water. The abundance of various salt components in these particles, as well as the inferred basic pH, exhibit a compelling similarity to the predicted composition of a subsurface Enceladus ocean in contact with its rock core[11]. The plume vapour is expected to be free of atomic sodium. Thus, the absence of sodium from optical spectra[12] is in good agreement with our results. In the E ring the upper limit for spectroscopy[12] is insufficiently sensitive to detect the concentrations we found.

The compositional *in situ* analysis of E-ring grains provides unique information about the composition of plume material combined with otherwise unachievable detection statistics. The Cosmic Dust Analyser[13] (CDA) aboard the Cassini spacecraft recorded thousands of mass spectra of E-ring particles, predominantly with radii of 0.1–1 μm. The instrument's chemical analyser subsystem produces time-of-flight mass spectra of cations and cationic aggregates of the gas and plasma cloud generated by high-velocity impacts of single grains onto a metal target.

Previous CDA results show that the bulk material of E-ring particles is water ice[9,14]. There appear to be distinct families of spectra, types I and II of which have already been introduced in previous work[9]. Type I spectra imply almost pure water ice; type II spectra exhibit significant amounts of organic and/or siliceous material. These non-water components alone imply the past or present interaction of liquid water with a rocky Enceladian core[9].

Sodium is considered the perfect indicator for a still existing liquid[11]. In the case of a slow downward freezing, as would be expected for icy planetary bodies, sodium salts previously dissolved from rock remain in the liquid phase, so the ice crust is expected to be almost salt-free[11]. Therefore, sublimation from heated water ice[15] or clathrate decomposition[16] in the crust cannot generate the measured Na concentrations. The detection of sodium in Enceladus' plume material is only plausible for a liquid source.

Approximately 93% of the E-ring spectra in the evaluated data set, obtained between October 2004 and December 2005 (Supplementary Information), show a Na^+ mass line and/or cluster (aggregate) ions containing Na^+ (Fig. 1). Cluster formation occurs during the impact and is very sensitive to the Na/H_2O ratio[17], providing a method of constraining the Na concentrations of the ice particles (Fig. 1).

About 6% of these spectra (hereafter 'type III' spectra) show a unique mass line pattern which we now attribute to ice particles with particularly high sodium contents (Fig. 1b). In addition to water and sodium they exhibit features that indicate the sodium salts NaCl, $NaHCO_3$ and/or Na_2CO_3 as significant constituents. Weaker K^+ mass lines suggest a minor contribution from potassium salts. On average, K^+ is 100 to 200 times less abundant than Na^+. The majority of E-ring ice spectra are of types I and II[9]. Na features are identified in over 90% of these spectra (a substantially higher percentage than found in the previous analysis, which considered only the most prominent mass lines[9]), albeit at much lower abundances than in the type III population (Fig. 1a).

In previous work, the Na^+ mass lines of Na-poor type I and II spectra were tentatively attributed to Na-bearing compounds having contaminated the detector's target before Cassini's launch[9,14]. However, the concentration of salts indicated by type III spectra is above any possible level of target contamination[18]. Furthermore, contamination as a dominant origin of the frequent low-abundance Na^+ mass lines in type I and II spectra can also be ruled out, because a careful analysis[18] significantly restricts the possible degree of Na contamination (Supplementary Information).

To constrain our interpretation further, we have recorded mass spectra of water aggregates generated from a laser-dispersed liquid[19] from a solution with different sodium salt concentrations (Supplementary Information). All Na-rich type III spectra mass lines (Fig. 1b) are reproduced well by laboratory spectra (Fig. 2b) from a $NaCl/NaHCO_3$ solution with 0.05–0.2 mol kg^{-1} NaCl and 2–5 times more NaCl than $NaHCO_3$. To reproduce the Na-poor E-ring spectra (Fig. 1a), very dilute salt solutions ($Na/H_2O < 10^{-7}$) are required (Fig. 2a). This implies at least two populations of ice grains, whose Na content differs by more than four orders of magnitude. Spectra with intermediate Na-concentrations are rare (<3%).

We suggest that the Na-rich grains are directly frozen submicrometre-sized droplets from the liquid plume reservoir. We employ a liquid dispersion model[19,20], assuming that—above the salt solution—aerosol-like droplets are produced (Fig. 3a), possibly by ascending bubbles of plume gases[4] (CO_2, N_2, CO, CH_4) in the liquid or by other processes that can disperse the liquid under the given conditions.

Another effective mechanism for grain formation is direct condensation of the ascending vapour within the ice channels[7]. The low Na content of type I and II grains (which amount to more than 90% of the recorded spectra) might thus directly reflect the low concentration of NaCl molecules in the water vapour above the liquid (Fig. 3b).

[1]Institut für Geowissenschaften, Universität Heidelberg, 69120 Heidelberg, Germany. [2]Max-Planck-Institut für Kernphysik, 69117 Heidelberg, Germany. [3]IGEP,Technische Universität Braunschweig, 38106 Braunschweig, Germany. [4]Nichtlineare Dynamik, Universität Potsdam, 14476 Potsdam–Golm, Germany. [5]Department of Mathematics, University of Leicester, Leicester LEI 7RH, UK. [6]Department of Physics, Moscow State University, 119991 Moscow, Russia. [7]Institut für Physikalische Chemie, Universität Göttingen, 37077 Göttingen, Germany. [8]Wilhelm-Oswald-Institut für Physikalische und Theoretische Chemie, Universität Leipzig, 04103 Leipzig, Germany. [9]Max-Planck-Institut für Dynamik und Selbstorganisation, 37073 Göttingen, Germany.

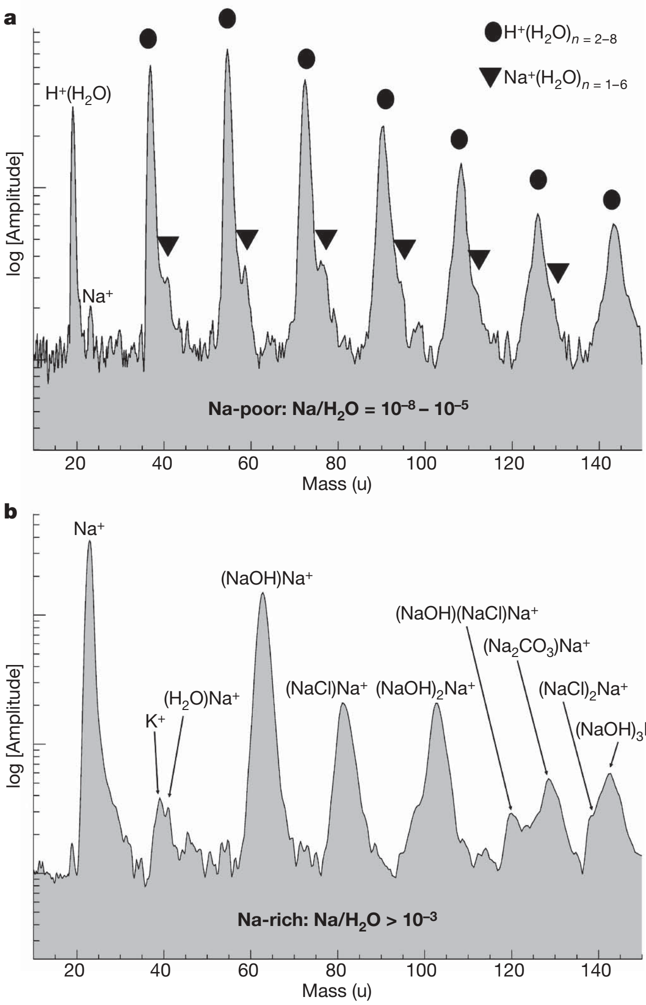


Figure 1 | Spectra of E-ring particles with differing Na/H_2O ratio. **a**, Spectrum of a Na-poor ice particle (types I and II), dominated by a sequence of water cluster ions of the form $H(H_2O)_n^+$, ($n = 2$–8). Weak mass lines of Na^+ and Na-hydrate ions ($Na(H_2O)_n^+$) are also present. The sensitivity of the detector to Na is at least several hundred times higher than to water, so the respective ice grains contain only traces of sodium. From the peak amplitudes we infer a varying Na/H_2O mixing ratio within the Na-poor ice grains (types I and II) of $5 \times 10^{-5(\pm 1)}$ to $5 \times 10^{-8(\pm 1)}$ (Supplementary Information). Often K^+ (39u, where u means unified atomic mass units, or daltons) and its respective hydrates—$K(H_2O)_n^+$—are also detectable. However, reliable determinations of the K concentrations have not been obtained, owing to overlapping of K spectral features with abundant water and Na-cluster mass lines. **b**, Co-added spectra of Na-rich water-ice particles (type III). Although the particles are still predominantly water, these spectra typically show very few pure water and Na-hydrate clusters, if any. They are characterized by an abundant Na^+ mass line followed by a peak sequence of hydroxyl-cluster-ions $Na(NaOH)_n^+$, representing the defining pattern of a type III spectrum. Laboratory experiments (Fig. 2) confirm that this cluster type is dominant only if the Na/H_2O ratio exceeds 10^{-3}. The pronounced appearance of $Na(NaOH)_n$ signatures also indicate an alkaline liquid. In addition, these spectra often show $Na(NaCl)_n^+$, (where $n = 1$–3), and $Na(Na_2CO_3)^+$, which implies that NaCl followed by $NaHCO_3$ and/or Na_2CO_3 are the main Na-bearing compounds (also Fig. 2 and Supplementary Information). A small amount of K^+ is also present. These are exactly the four most-abundant non-water species predicted to exist in an ocean within Enceladus[11].

Thermodynamically, a highly depleted NaCl gas phase in equilibrium with non-dissociated NaCl in the liquid[11] is in fact expected (Fig. 4 and Supplementary Information). A fraction of Na-free type I grains might also be ejecta from the surface of Enceladus, caused by micro-meteoroid impacts[9].

For an ocean in equilibrium with a rocky Enceladus core, Na^+ concentrations of 0.05–0.1 $mol\,kg^{-1}$, an overall salinity of

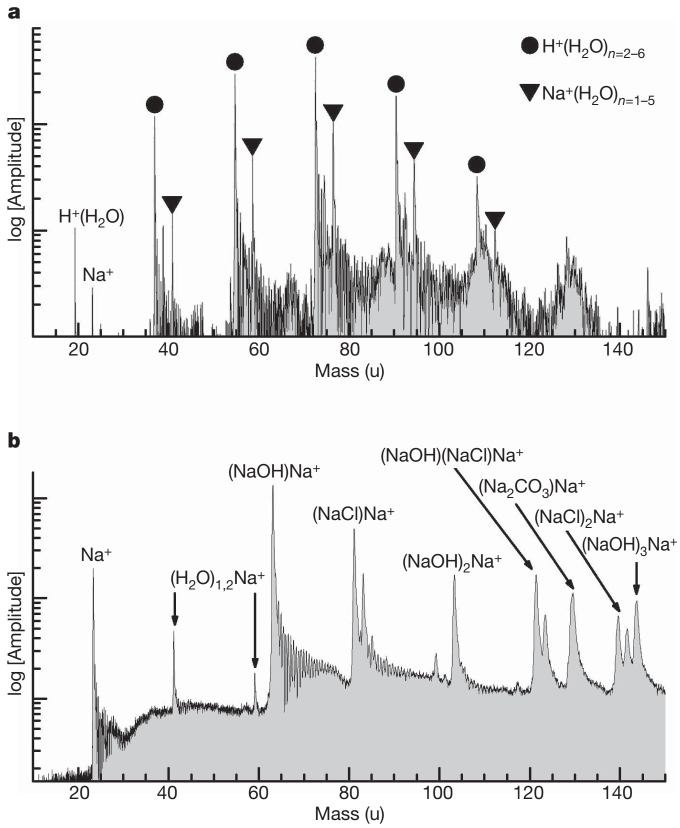


Figure 2 | Reproduction of CDA measurements in the laboratory by laser dispersion[19] of salt water. Differences between the laboratory experiments (laser shot on a liquid) and CDA (high-speed impact of an ice grain), have no effect on our interpretation (Supplementary Information). What is important is that both measurements, in the laboratory and at Enceladus, start from a similar liquid solution. **a**, Cation spectrum of water with $10^{-6}\,mol\,kg^{-1}$ NaCl ($Na/H_2O \approx 2 \times 10^{-8}$). It shows the same cluster characteristics as Na-poor (types I and II) E-ring grains (Fig. 1a). $H(H_2O)_n^+$ clusters dominate but Na^+ and its hydrates—$Na(H_2O)_n^+$—are also present. The NaCl concentration of Na-poor spectra inferred from the laboratory experiments is at the lower boundary of our earlier estimates (Fig. 1a). **b**, Ion spectrum of water with 0.2 $mol\,kg^{-1}$ NaCl and 0.1 $mol\,kg^{-1}$ $NaHCO_3$ ($Na/H_2O \approx 5 \times 10^{-3}$), pH = 9. Except for the better resolution, all characteristic mass lines are identical to a Na-rich (type III) E-ring grain spectrum (Fig. 1b): few Na-hydrates but abundant $Na(NaOH)_n^+$, $Na(NaCl)_n^+$ and $Na(Na_2CO_3)^+$ clusters are observed. The cluster pattern not only allows for estimates of the NaCl, $NaHCO_3$ and KCl content, but also of the pH of a liquid solution with the ice grain composition.

3–8 $g\,kg^{-1}$, and an alkaline pH of 8–11 have been predicted[11]. The increasing proportion of salts within the remaining liquid phase, caused by slow downward freezing of an early ocean, has also been quantified[11]. Using the correlation between laboratory experiments and CDA spectra, we find that the composition of type III E-ring grains agrees with the modelled Enceladus ocean composition. Our values for the concentrations of NaCl (0.05–0.2 $mol\,kg^{-1}$) and the Na-carbonates (0.02–0.1 $mol\,kg^{-1}$) as well as the inferred alkalinity (pH 8.5–9) match the model. However, the Na/K ratio is about ten times higher than predicted[11]. This may be due to K-deposition in clays at the rock-water interface, a process well known from terrestrial waters[21]. Our measurements imply a 'soda-ocean' rich in bicarbonate and/or carbonate. This is consistent with the abundant CO_2 observed in the plume vapour[4]. Escaping gases[4] are probably also generated by clathrate decomposition at the water–ice boundary or hydrothermal processes[10]. Alkaline salt water, together with the observed organic compounds[4,9] and the thermal energy obviously present in the south polar region[6], could provide an environment well suited for the formation of life precursors[22,23].

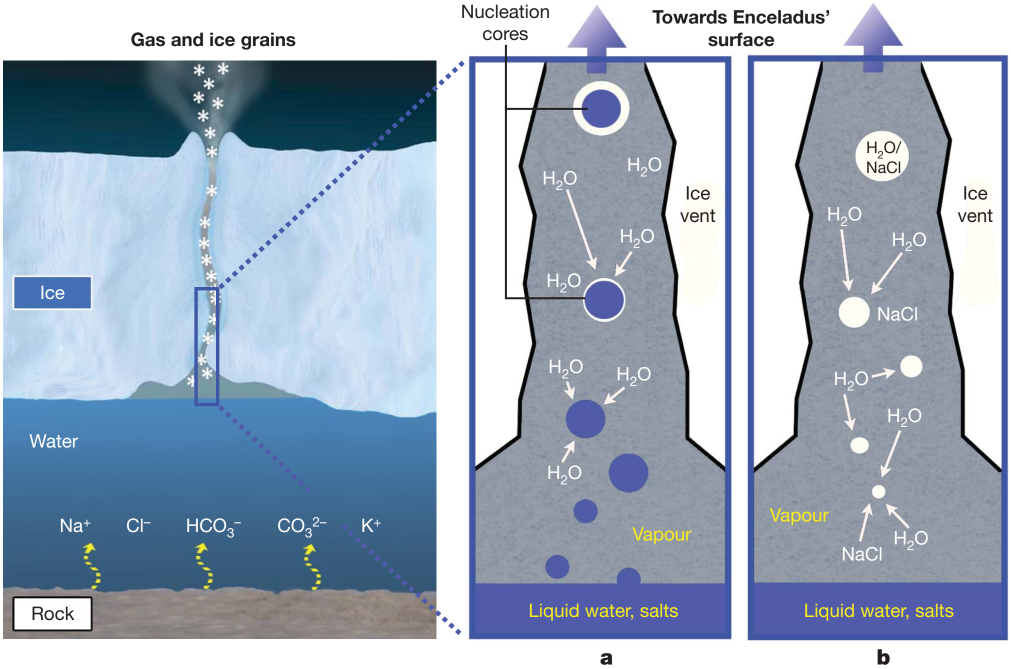


Figure 3 | Schematic of how liquid dispersion and condensation from vapour form ice particles with very different Na content. a, Na-rich grains. Aerosol-like droplets (blue) are generated by dispersion above the salty liquid. They preserve the ocean's composition. Such droplets (submicrometre- to micrometre-sized) would condense some additional water from the supersaturated gas[7] that drags them towards the surface. Therefore, the salt content of Na-rich (type III) grains can be considered a lower limit for the surface of an Enceladus ocean. **b**, Na-poor grains. The standard formation of ice grains by condensation of water vapour ascending from the liquid reservoir with subsequent growth and transport in the vent[7]. The particles initially have the same low ratio of Na/H_2O as the gas (Fig. 4). By direct or indirect interaction (such as via crack walls) with the Na-rich grains, they may be further enriched with Na-salt traces. Therefore, the salt content of the Na-poor grains can be considered an upper limit for that of the plume gas.

Individual plume sources stay active for years[24], implying outflow from a large reservoir. Moreover, given the observed gas production rate[2] ($\sim$150–300 kg s^{-1}), it can be shown (using heat-flow arguments) that the area of the liquid–gas interface must be at least of the order of square kilometres (Supplementary Information). This requires large vapour chambers above the liquid that narrow to the vent channels. A local enhancement of salinity (owing to evaporation) will be rapidly mixed with the bulk of the liquid. The salt-enrichment of the reservoir is a slow process, not violating the steady-state assumption for vent dynamics[7]. For example, the timescale for evaporation (and increasing salinity) even for an isolated liquid reservoir of an equivalent radius of 10 km would be 300,000 years (Supplementary Information).

So far, no atomic Na or K has been detected in Enceladus' environment by spectroscopy[12]. This does not contradict our results. The sodium in the plumes is almost entirely located in the type III grains. On a very low level, dissolved alkali salts (NaCl and KCl) evaporate only in molecular (not atomic) form (Fig. 4 and Supplementary Information). Moreover, any free alkali atoms will quickly react in a water vapour environment[25–27]. However, particles are subject to plasma sputtering in the E ring and over decades embedded alkali compounds will be slowly released from the grains. The current detection limit for optical Na emission in E-ring gas is equivalent to a Na/H_2O mixing ratio[12] of 7×10^{-6}. Making the (simplified) assumption of similar particle masses in Na-rich and Na-poor populations, the average Na/H_2O ratio (types I, II and III) in E-ring ice grains is estimated to be 3×10^{-4} to 3×10^{-5}. However, most of the total ejected grain mass falls back to the moon's surface[7,28] (the cryovolcanically active regions should be covered with Na-bearing plume particles), so the particle total mass ejected into the E ring is about 200 times less than the mass of water vapour[2,5]. Therefore, the overall Na/H_2O E-ring mixing ratio is diluted to roughly 10^{-6}–10^{-7}, significantly below the optical detection limit.

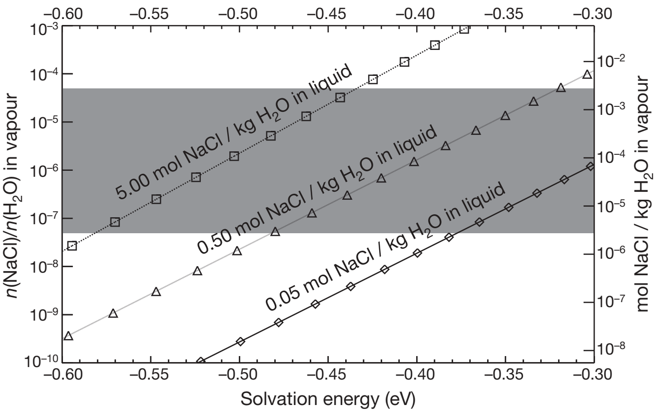


Figure 4 | Equilibrium concentrations of NaCl in the vapour phase above a subsurface Enceladus water reservoir. The calculated NaCl vapour concentration (Supplementary Information) is plotted for the case of a slightly salty ocean (0.05 mol kg^{-1}, diamonds), an earth-like ocean (0.5 mol kg^{-1}, triangles) and a NaCl brine close to the eutectic concentration (5 mol kg^{-1}, squares). The dissolved ions, Na^+ and Cl^-, cannot evaporate owing to their high solvation energy, so the concentration of NaCl in the vapour depends on the solvation energy of the small fraction of undissociated NaCl dipoles existing in the liquid. We constrained the range of plausible solvation energies to 0.3–0.6 eV per molecule (Supplementary Information). The grey area shows the Na concentration inferred for Na-poor (types I and II) E-ring particles. For a wide range of plausible solvation energies and NaCl molalities in the ocean, the measured Na content is in agreement with that of vapour evaporated from a salty ocean. The initially low vapour concentration of molecular alkali compounds is probably further depleted by the higher condensation rates of such refractory compounds (for example, at the crack walls).

Received 4 September 2008; accepted 6 April 2009.

1. Dougherty, M. K. *et al.* Identification of a dynamic atmosphere at Enceladus with the Cassini magnetometer. *Science* **311**, 1406–1409 (2006).
2. Hansen, C. J. *et al.* Enceladus' water vapor plume. *Science* **311**, 1422–1425 (2006).
3. Spahn, F. *et al.* Cassini dust measurements at Enceladus and implications for the origin of the E ring. *Science* **311**, 1416–1418 (2006).

4. Waite, J. H. *et al.* Cassini ion and neutral mass spectrometer: Enceladus plume composition and structure. *Science* **311,** 1419–1422 (2006).
5. Porco, C. C. *et al.* Cassini observes the active south pole of Enceladus. *Science* **311,** 1393–1401 (2006).
6. Spencer, J. R. *et al.* Cassini encounters Enceladus: background and the discovery of a south polar hot spot. *Science* **311,** 1401–1405 (2006).
7. Schmidt, J., Brilliantov, N., Spahn, F. & Kempf, S. Slow dust in Enceladus' plume from condensation and wall collisions in tiger stripe fractures. *Nature* **451,** 685–688 (2008).
8. Kempf, S. *et al.* The E ring in the vicinity of Enceladus. I. Spatial distribution and properties of the ring particles. *Icarus* **193,** 420–437 (2008).
9. Postberg, F. *et al.* The E ring in the vicinity of Enceladus. II. Probing the moon's interior—the composition of E-ring particles. *Icarus* **193,** 438–454 (2008).
10. Matson, D. L., Castillo, J. C., Lunine, J. & Johnson, T. V. Enceladus' plume: compositional evidence for a hot interior. *Icarus* **187,** 569–573 (2007).
11. Zolotov, M. Y. An oceanic composition on early and today's Enceladus. *Geophys. Res. Lett.* **34,** L23203 (2007).
12. Schneider, N. M. *et al.* No sodium in the vapour plumes of Enceladus. *Nature* doi:10.1038/nature08070 (this issue).
13. Srama, R. *et al.* The Cassini cosmic dust analyzer. *Space Sci. Rev.* **114,** 465–518 (2004).
14. Hillier, J. K. *et al.* The composition of Saturn's E ring. *Mon. Not. R. Astron. Soc.* **388,** 1588–1596 (2007).
15. Nimmo, F., Spencer, J. R., Papparlado, R. T. & Mullen, M. E. Shear heating as the origin of the plumes and heat flux on Enceladus. *Nature* **447,** 289–291 (2007).
16. Kieffer, S. W. *et al.* A clathrate reservoir hypothesis for Enceladus' south polar plume. *Science* **314,** 1764–1766 (2006).
17. Steinbach, C. & Buck, U. Reaction and solvation of sodium in hydrogen bonded solvent clusters. *Phys. Chem. Chem. Phys.* **7,** 986–990 (2005).
18. Postberg, F. *et al.* Discriminating contamination from particle components in spectra of Cassini's dust detector CDA. *Planet. Space Sci.* (in the press).
19. Charvat, A. & Abel, B. How to make big molecules fly out of liquid water: applications, features and physics of laser assisted liquid phase dispersion mass spectrometry. *Phys. Chem. Chem. Phys.* **9,** 3335–3360 (2007).
20. Frohn, A. & Roth, N. *Dynamics of Droplets* 245–260 (Springer, 2000).
21. Harder, H. Über das Kalium-Natrium-Verhältnis in Gewässern und die Tonmineralbildung. *Naturwissenschaften* **54,** 613 (1967).
22. Kempe, S. & Kazmierczak, J. Biogenesis and early life on Earth and Europa: favoured by an alkaline ocean? *Astrobiology* **2,** 123–130 (2002).
23. McKay, C. P., Porco, C. C., Altheide, T., Davis, W. L. & Kral, T. A. The possible origin and persistence of life on Enceladus and detection of biomarkers in the plume. *Astrobiology* **8,** 909–919 (2008).
24. Spitale, J. N. & Porco, C. C. Association of jets on Enceladus with the warmest regions on its south-polar fractures. *Nature* **449,** 695–697 (2007).
25. Dzidic, I. & Kebarle, P. Hydration of the alkali ions in the gas phase. *J. Phys. Chem.* **74,** 1466–1474 (1969).
26. Schulz, C. P., Haugstätter, R., Tittes, H. U. & Hertel, I. V. Free sodium-water clusters. *Phys. Rev. Lett.* **57,** 1703–1706 (1986).
27. Buck, U. & Steinbach, C. Formation of sodium hydroxide in multiple sodium-water cluster collisions. *J. Phys. Chem. A* **102,** 7333–7336 (1998).
28. Kempf, S., Beckmann, U. & Schmidt, J. How the Enceladus dust plumes form Saturn's E ring. *Icarus*. (submitted).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank U. Beckmann, M. Burger, M. Burton, M. Gellert, E. Grün, J. K. Hillier, N. Schneider, F. Spahn, and M. Zolotov for discussions. We acknowledge the efforts of the Cassini team and JPL. The work has been supported by the DLR, the DFG and the Frontier programme of the University of Heidelberg.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to F.P. (frank.postberg@mpi-hd.mpg.de).

nature Vol 459|25 June 2009|doi:10.1038/nature08070

LETTERS

No sodium in the vapour plumes of Enceladus

Nicholas M. Schneider[1], Matthew H. Burger[2], Emily L. Schaller[3], Michael E. Brown[4], Robert E. Johnson[5], Jeffrey S. Kargel[6], Michele K. Dougherty[7] & Nicholas A. Achilleos[8]

The discovery of water vapour and ice particles erupting from Saturn's moon Enceladus fuelled speculation that an internal ocean was the source[1–3]. Alternatively, the source might be ice warmed, melted or crushed by tectonic motions[4]. Sodium chloride (that is, salt) is expected to be present in a long-lived ocean in contact with a rocky core. Here we report a ground-based spectroscopic search for atomic sodium near Enceladus that places an upper limit on the mixing ratio in the vapour plumes orders of magnitude below the expected ocean salinity[5]. The low sodium content of escaping vapour, together with the small fraction of salt-bearing particles[6], argues against a situation in which a near-surface geyser is fuelled by a salty ocean through cracks in the crust[1]. The lack of observable sodium in the vapour is consistent with a wide variety of alternative eruption sources, including a deep ocean[6], a freshwater reservoir, or ice. The existing data may be insufficient to distinguish between these hypotheses.

The discovery of icy particles containing ~1% sodium chloride in Saturn's E ring has led to the suggestion that they formed from a salty ocean below Enceladus' icy crust[6]. But the vast majority of escaping matter is gaseous—mass proportions for vapour:low-sodium particles:1%-sodium particles are[6] approximately 4,000:20:1. It is therefore essential to examine the bulk of material escaping Enceladus to test and further constrain the hypothesis. If sodium is present in atomic form in the vapour, it will be readily visible in the plume near Enceladus; if sodium is ejected in molecular form, it will photodissociate into atomic form and will orbit Saturn for weeks to months and again be readily observed. Escaping sodium is a surprisingly common phenomenon in the Solar System, being easily detected at Io[7], Europa[8], Mercury[9], the Moon[10] and comets[11], and sodium chloride vapour has been detected above Io's volcanoes[12]

Sodium is arguably the most easily detected atom among common Solar System elements. Its atomic structure places its allowed transitions near the peak of the solar spectrum, so even small amounts of sodium appear bright in resonant scattering of sunlight. We undertook a deep search for atomic sodium D-line emission (589.0, 589.6 nm) near Enceladus and Saturn's E ring using high-resolution long-slit spectroscopy. We used the 10-m Keck 1 telescope with the HIRES spectrograph[13] and the 3.9-m Anglo-Australian Telescope (AAT) with the UCLES spectrograph[14] (Table 1). The high resolution of echelle spectrographs disperses the solar continuum reflected by Enceladus, as well as the solar continuum from Saturn and its rings scattered by Earth's atmosphere and the telescope optics into the spectrograph slit. The long slit allows the determination of independent spectra from many locations along the slit that can be summed to increase signal-to-noise ratio or differenced for sky background subtraction.

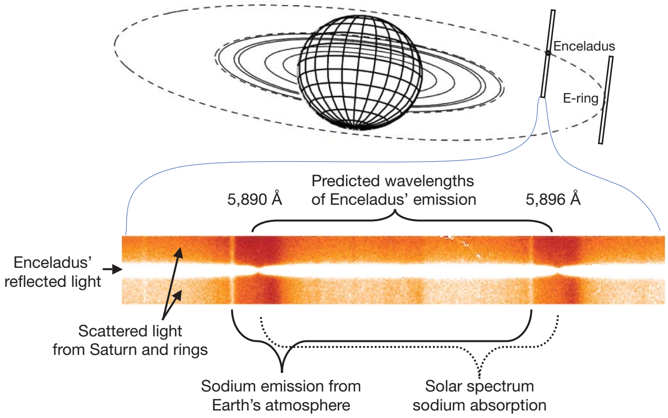


Figure 1 | Observing geometry and corresponding sodium D-line spectrum of Enceladus and nearby E ring. Bottom, 30-min spatially resolved spectrum obtained with the Keck 1 HIRES spectrograph on 25 February 2007 at 09:59 UT, with wavelength information displayed horizontally and spatial information displayed vertically. The Enceladus and scattered Saturn/ring spectra display the strong solar Fraunhofer D-line absorptions. Sodium features appear at different wavelengths due to their respective Doppler shifts; the predicted Doppler-shifted locations of emission from Enceladus sodium are indicated. Top, diagram showing geometry of Saturn, its E ring (outer dashed line), slit position for the spectrum shown, and also the observing geometry for dedicated E-ring observations. The spatial resolution of the Keck observations is 0.8″ full-width at half-maximum, corresponding to ~5,000 km or ~20 Enceladus radii.

Figure 1 shows the observing geometry and a typical Keck spectral image. Table 1 summarizes the brightness upper limits obtained from both telescopes and derived sodium abundances (Supplementary Information). The Keck observations place more restrictive upper limits owing to this telescope's larger aperture and lower scattered

Table 1 | Observations

Observatory/ instrument	Dates (UT)	Resolving power	Slit size (arcsec)	Exposure times (min)	No. of expo-sures	Upper limit near Enceladus		Upper limit near E ring	
						Brightness (R)	Column abundance (atoms cm^{-2})	Brightness (R)	Column abundance (atoms cm^{-2})
Keck 1/ HIRES	25 Feb. 2007	60,000	14×0.57	30	13	4–8	$(0.5–1.0) \times 10^8$	2	2.5×10^7
AAT/ UCLES	27 Feb. –2 Mar. 2007	80,000	12×0.5	20	21			100	1.3×10^9

In the seventh and eighth columns are shown the upper limit of sodium D-line emission brightness, and the derived sodium column abundance. AAT, Anglo-Australian Telescope.

[1]Laboratory for Atmospheric & Space Physics, University of Colorado, Boulder, Colorado 80309, USA. [2]University of Maryland, Baltimore County, and NASA/Goddard Space Flight Center, Greenbelt, Maryland 20771, USA. [3]Institute for Astronomy, University of Hawaii, Honolulu, Hawaii 96822, USA. [4]Geological & Planetary Sciences, Caltech, Pasadena, California 91125, USA. [5]Engineering Physics, University of Virginia, Charlottesville, Virginia 22904, USA. [6]Hydrology & Water Resources, University of Arizona, Tucson, Arizona 85721, USA. [7]Space & Atmospheric Physics, Imperial College, London SW7 2AZ, UK. [8]Physics & Astronomy, University College, London WC1E 6BT, UK.

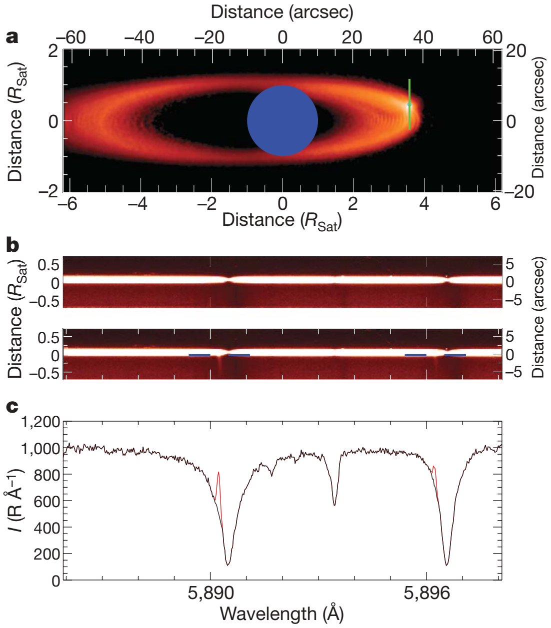


Figure 2 | Determining the 4 × 10^{-7} upper limit for direct sodium ejection into a south-directed plume. **a**, Plume model image for the geometry of Keck Enceladus observations. Intensity scaling is logarithmic. The model uses an atomic sodium escape rate of 0.4 g s^{-1}, corresponding to an Na/H_2O mixing ratio of 2×10^{-6}. The green line represents the slit across Enceladus. The asymmetry is caused by solar radiation pressure on atomic sodium. R_{Sat}, Saturn radius. **b**, Sum of all sky-subtracted spectral images for Enceladus observations with (bottom) and without (top) 40 R of simulated sodium plume emission. The simulated signal does not include photon noise. **c**, Spectra of the spatial region indicated by the blue lines in **b** located along the lower edge of Enceladus' continuum near the simulated emissions. The black curve is data alone, the red curve adds 40 R of simulated emission derived from the model in **a**. This simulated emission would be readily detectable; the actual upper limit derived from the original spectrum is 8 R, yielding Na/$H_2O < 4 \times 10^{-7}$. Further information on the derivation of the upper limits is given in the Supplementary Information. I, intensity.

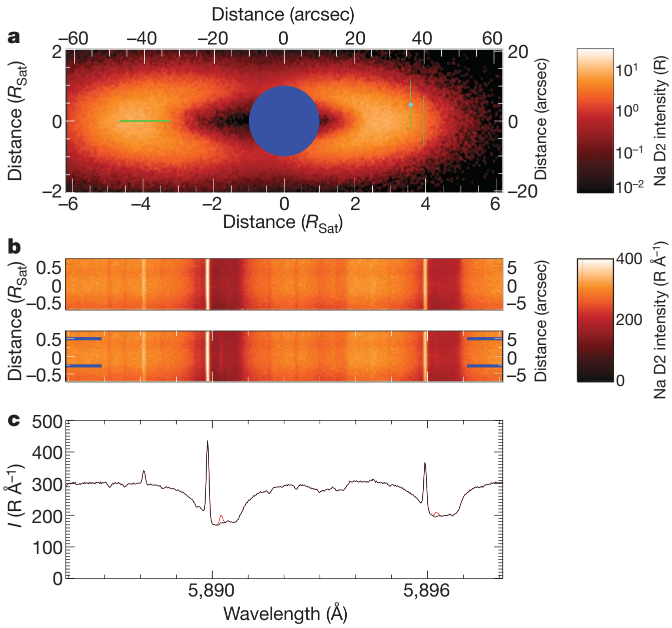


Figure 3 | Determining the 7 × 10^{-6} upper limit for a molecular source, with panels displaying the same information as in Fig. 2. **a**, Model image for sodium produced by molecular dissociation as described in the text. The model uses an atomic sodium escape rate of 2 g s^{-1}, corresponding to an Na/H_2O mixing ratio of 1×10^{-5}. The green lines represent, from left to right, the AAT east E-ring ansa slit, the Keck slit at Enceladus, and the Keck slit at the west E-ring ansa. (An ansa is the outer edge of a ring as projected on the sky.) The spatial distribution for sodium sputtered from E-ring particles is essentially the same. **b**, Sum of all spectral images for Keck E-ring observations, with and without 3 R of simulated emission. **c**, Spectra summed over the spatial region indicated in **b** by the horizontal blue lines at the left and right edges. The black curve shows data alone; the red curve adds 3 R of simulated E-ring sodium emission. The observed upper limit from these data is 2 R, yielding Na/$H_2O < 7 \times 10^{-6}$. Similar analysis of the Keck Enceladus slit positions yield Na/$H_2O < 1.4 \times 10^{-5}$, and the AAT observations find Na/$H_2O < 4 \times 10^{-4}$. The upper limit for molecular ejection is higher than for atomic ejection, as sodium is first dispersed in undetectable molecular form far from Enceladus into a ring around Saturn, then dispersed from that ring through dissociation, thereby filling a much larger volume. The Fraunhofer lines are broader than those in Fig. 2 owing to the roughly equal admixture of scattered light from Saturn and from the rings at different Doppler shifts.

light background. In summed Keck E-ring spectral images (Figs 2b, 3b) we find an upper limit of 2 Rayleighs (2 R) on D-line emission. In the Keck Enceladus spectral image (Fig. 2b), where stray light from the satellite increases the background, we place an upper limit of 4 R at distances ≥2″ from Enceladus and 8 R at 1″. Each spectral image contains three slit-filling night sky emissions of brightnesses 4 R, 32 R, 17 R from left to right. (The first line comes from OH and the other two from sodium; they appear at approximate brightnesses and relative strengths consistent with past observations[15].) The ease of detecting the 4 R airglow line confirms we can confidently rule out emissions of this magnitude from Enceladus, the E ring or even a cloud larger than the slit length. For reference, the observed sodium brightnesses near Io and Europa are 200 kR and 10 kR; sodium emission from Io was detectable on these nights within seconds but could not be detected at Enceladus in more than six hours.

We convert the limiting column abundance into an upper limit on the plume sodium content by ratioing a modelled upper limit on sodium escape to the 150 kg s^{-1} of H_2O escape[16–18]. Our Monte Carlo model includes all major source, loss and transport processes for atomic sodium, and has been used to study neutral clouds at Io[19], Europa[20] and Enceladus[21] (Supplementary Information).

Figures 2 and 3 also show model results for sodium ejection in atomic or molecular form. With atomic ejection, sodium is visible immediately on ejection. Sodium eventually forms a complete torus around Saturn, but the high concentration in the plume near Enceladus dominates the brightness. The 8 R upper limit at the plume position near Enceladus translates to a very stringent upper limit of Na/$H_2O < 4 \times 10^{-7}$. Sodium escaping in molecular form (Fig. 3) does not become visible until photodissociation after forming a narrow molecular torus around Saturn. Keck E-ring slit pointings constrain Na/$H_2O < 7 \times 10^{-6}$, and the AAT data independently constrain Na/$H_2O < 4 \times 10^{-4}$. Sodium supplied by E-ring particle sputtering will resemble Fig. 3. The sodium mixing ratio of these particles[6] (referenced to the observed H_2O escape) is $\sim 5 \times 10^{-7}$, so our Keck upper limits are consistent with the detection reported in the companion paper[6].

Our observations show that the bulk of mass escaping Enceladus has much lower salt content than ocean models. Our upper limit for atomic escape, Na/$H_2O < 4 \times 10^{-7}$, is 5,000 times smaller than the mixing ratio (2×10^{-3}) of the sodium-rich ocean model[5], and our molecular upper limit, Na/$H_2O < 7 \times 10^{-6}$, is still 30 times lower than the 2×10^{-4} sodium-poor ocean model[5]. In fact, the lack of detectable sodium makes Enceladus unusual among objects with thin atmospheres and measurable atmospheric escape, namely Io, Europa, the Moon and Mercury.

The observed low sodium content in the vapour leads to five conclusions about plausible plume sources. First, the vapour could clearly result from evaporation from a low-salinity liquid reservoir, from warm ice sublimation, or from clathrate decomposition[22], though these mechanisms currently do not address the observed salt-bearing particles[6]. Second, the escaping gas cannot come from complete vaporization of ocean water including its dissolved salts, as might

occur from a near-surface geyser boiling into a vacuum[1]. Third, vapour low in sodium can originate from a deep salty ocean, but only if evaporation occurs in equilibrium at the interface of a large liquid/vapour reservoir, as detailed in the companion paper[6] and in ref. 23. Fourth, similar scenarios in which ocean water flows nearer the surface along narrow cracks may prove unworkable. In the confined vent plumbing, preferential evaporation of the more volatile H_2O could enrich salt content in the liquid, possibly choking off the vent or leading to the ejection of very high salinity particles, which are not observed[6]. Fifth, a hybrid scenario, in which a low-salinity reservoir becomes enriched in the vent plumbing, could explain both the low-sodium vapour and the observed salt-bearing particles. In this scenario, however, the low values of salinity would indicate little exposure to the rocky core. We conclude that the observed sodium content of particles and vapour argue against certain plume eruption models and constrain others, but are insufficient to distinguish between these remaining scenarios. It remains plausible that more than one mechanism is at work; models based on evaporation of water require a separate mechanism such as clathrate decomposition to explain other escaping gases observed by Cassini's INMS instrument[24].

Received 24 October 2008; accepted 8 April 2009.

1. Porco, C. C. *et al.* Cassini observes the active south pole of Enceladus. *Science* **311,** 1393–1401 (2006).
2. Spencer, J. R. *et al.* Cassini encounters Enceladus: background and the discovery of a south polar hot spot. *Science* **311,** 1401–1405 (2006).
3. McKay, C. P., Porco, C. C., Altheide, T., Davis, W. L. & Kral, T. A. The possible origin and persistence of life on Enceladus and detection of biomarkers in the plume. *Astrobiology* **8,** 909–919 (2008).
4. Nimmo, F., Spencer, J. R., Pappalardo, R. T. & Mullen, M. E. Shear heating as the origin of the plumes and heat flux on Enceladus. *Nature* **447,** 289–291 (2007).
5. Zolotov, M. Y. An oceanic composition on early and today's Enceladus. *Geophys. Res. Lett.* **34,** L23203, doi:10.1029/2007GL031234 (2007).
6. Postberg, F. *et al.* Sodium salts in E-ring ice grains from an ocean below the surface of Enceladus. *Nature* doi:10.1038/nature08046 (this issue).
7. Brown, R. A. in *Exploration of the Planetary System* (eds Woszczyk, A. & Iwaniszewska, C.) 527–531 (Proc. IAU Symp. No. 65, Reidel, 1974).
8. Brown, M. E. & Hill, R. E. Discovery of an extended sodium atmosphere around Europa. *Nature* **380,** 229–231 (1996).
9. Potter, A. E. & Morgan, T. H. Discovery of sodium in the atmosphere of Mercury. *Science* **229,** 651–653 (1985).
10. Potter, A. E. & Morgan, T. H. Discovery of sodium and potassium vapor in the atmosphere of the Moon. *Science* **241,** 675–680 (1988).
11. Levin, B. J. On the reported Na tails of comets. *Icarus* **4,** 497–499 (1964).
12. Lellouch, E., Paubert, G., Moses, J. I., Schneider, N. M. & Strobel, D. F. Volcanically emitted sodium chloride as a source for Io's neutral clouds and plasma torus. *Nature* **421,** 45–47 (2003).
13. Vogt, S. S. *et al.* HIRES: the high resolution echelle spectrometer on the Keck 10-m telescope. *Proc. SPIE* **2198,** 362–375 (1994).
14. Diego, F., Charalambous, A., Fish, A. C. & Walker, D. D. Final tests and commissioning of the UCL echelle spectrograph. *Proc. SPIE* **1235,** 562–576 (1990).
15. Osterbrock, D. E. *et al.* Night-sky high-resolution spectral atlas of OH and O_2 emission lines for echelle spectrograph wavelength calibration. *Publ. Astron. Soc. Pacif.* **108,** 277–308 (1996).
16. Hansen, C. J. *et al.* Enceladus' water vapor plume. *Science* **311,** 1422–1425 (2006).
17. Tian, F. *et al.* Monte Carlo simulations of the water vapor plumes on Enceladus. *Icarus* **188,** 154–161 (2007).
18. Jurac, S. *et al.* Saturn: search for a missing water source. *Geophys. Res. Lett.* **29,** 2172, doi:10.1029/2002GL015855 (2002).
19. Burger, M. H. *Io's Neutral Clouds: From the Atmosphere to the Plasma Torus.* Ph.D. thesis, Univ. Colorado (2003).
20. Burger, M. H. & Johnson, R. E. Europa's neutral cloud: morphology and comparisons to Io. *Icarus* **171,** 557–560 (2004).
21. Burger, M. H. *et al.* Understanding the escape of water from Enceladus. *J. Geophys. Res.* **112,** A06219, doi:10.1029/2002GL015855 (2007).
22. Kieffer, S. W. *et al.* A clathrate reservoir hypothesis for Enceladus' south polar plume. *Science* **314,** 1764–1766 (2006).
23. Schmidt, J., Brillantov, N., Spahn, F. & Kempf, S. Slow dust in Enceladus' plume from condensation and wall collisions in tiger stripe fractures. *Nature* **451,** 685–688 (2008).
24. Hunter Waite, J. Jr *et al.* Cassini ion and neutral mass spectrometer: Enceladus plume composition and structure. *Science* **311,** 1419–1422 (2006).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements Some of the data presented here were obtained at the W.M. Keck Observatory, which is operated as a scientific partnership among the California Institute of Technology, the University of California and the National Aeronautics and Space Administration. The Observatory was made possible by the financial support of the W.M. Keck Foundation. We also acknowledge the Anglo-Australian Telescope and its staff. This work was supported by the National Science Foundation's Planetary Astronomy Program and the NASA Postdoctoral Program. This paper has benefited from discussions with M. Zolotov, J. Spencer, C. Porco, T. Johnson, A. Ingersoll, W. McKinnon, C. Mackay, F. Postberg, J. Schmidt, S. Kempf and R. Pappalardo.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to N.S. (nick.schneider@lasp.colorado.edu).

LETTERS

Optically controlled locking of the nuclear field via coherent dark-state spectroscopy

Xiaodong Xu[1]*, Wang Yao[4]*, Bo Sun[1]*, Duncan G. Steel[1], Allan S. Bracker[2], Daniel Gammon[2] & L. J. Sham[3]

A single electron or hole spin trapped inside a semiconductor quantum dot forms the foundation for many proposed quantum logic devices[1–6]. In group III–V materials, the resonance and coherence between two ground states of the single spin are inevitably affected by the lattice nuclear spins through the hyperfine interaction[7–9], while the dynamics of the single spin also influence the nuclear environment[10–15]. Recent efforts[12,16] have been made to protect the coherence of spins in quantum dots by suppressing the nuclear spin fluctuations. However, coherent control of a single spin in a single dot with simultaneous suppression of the nuclear fluctuations has yet to be achieved. Here we report the suppression of nuclear field fluctuations in a singly charged quantum dot to well below the thermal value, as shown by an enhancement of the single electron spin dephasing time T_2*, which we measure using coherent dark-state spectroscopy. The suppression of nuclear fluctuations is found to result from a hole-spin assisted dynamic nuclear spin polarization feedback process, where the stable value of the nuclear field is determined only by the laser frequencies at fixed laser powers. This nuclear field locking is further demonstrated in a three-laser measurement, indicating a possible enhancement of the electron spin T_2* by a factor of several hundred. This is a simple and powerful method of enhancing the electron spin coherence time without use of 'spin echo'-type techniques[8,12]. We expect that our results will enable the reproducible preparation of the nuclear spin environment for repetitive control and measurement of a single spin with minimal statistical broadening.

We performed the experiment on a single negatively charged quantum dot embedded in a Schottky diode structure. Figure 1a shows the four-level energy diagram of the trion states under an external magnetic field perpendicular to the sample growth direction. In the pump–probe experiment, two narrow-linewidth continuous wave lasers selectively excite a three-level lambda subsystem, as shown in the dashed box of Fig. 1a (ref. 17; see Methods Summary). When the pump and probe lasers exactly match the two-photon Raman resonance (TPR) condition, a coherent superposition of the spin ground states are formed[18]. This is known as the dark state[19] and represents a coherent manipulation of the electron spin in the frequency domain (a brief comparison of ref. 18 with the

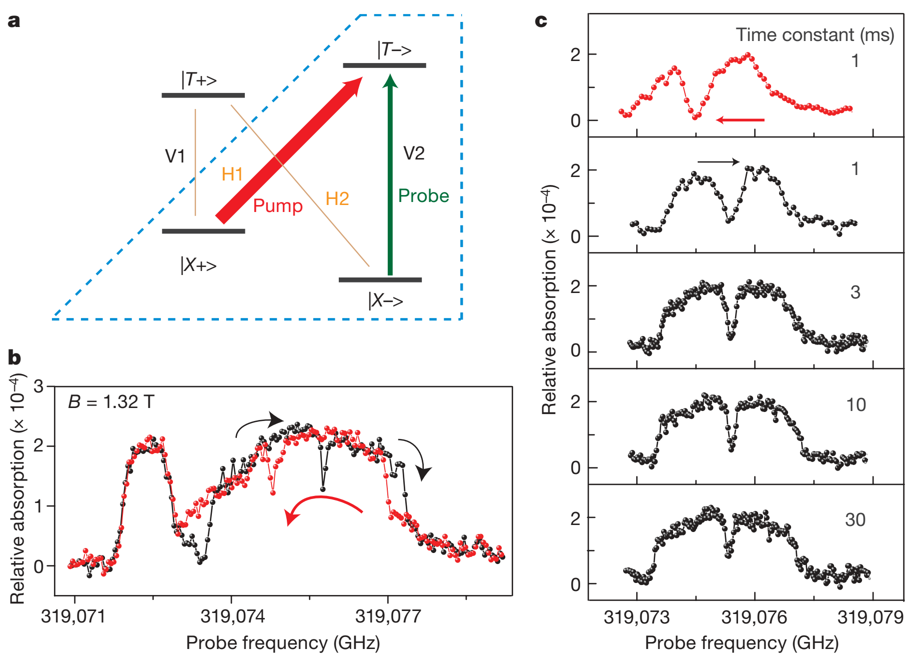


Figure 1 | Laser frequency sweep direction dependent probe absorption spectrum. a, The trion energy level diagram with a magnetic field applied in the Voigt geometry. The blue dashed box indicates the selected three-level Lambda system. A strong pump beam is near-resonant with transition H1 and a weak beam probes transition V2. **b,** The probe absorption spectrum at an external magnetic field of 1.32 T. The black (or red) curve represents the probe absorption spectrum of the forward (or backward) scan. **c,** The probe absorption spectrum as a function of the laser scan rate, indicated by the lock-in time constant. The top red curve is the backward scan with a 1-ms lock-in time constant.

[1]The H. M. Randall Laboratory of Physics, The University of Michigan, Ann Arbor, Michigan 48109, USA. [2]Naval Research Laboratory, Washington DC 20375, USA. [3]Department of Physics, University of California San Diego, La Jolla, California 92093, USA. [4]Department of Physics, The University of Hong Kong, Hong Kong, China.
*These authors contributed equally to this work.

current experiment can be found in the Supplementary Information). Because absorption of the probe laser by trion excitation increases abruptly with detuning from the TPR, the generated dark state is very sensitive to small changes in the nuclear field, which can affect the TPR and hence can be used as an *in situ* probe of the nuclear spin environment in a quantum dot[20].

Figure 1b shows the probe absorption spectrum at a magnetic field of 1.32 T. The black curve is the spectrum obtained by sweeping the probe laser from low frequency to high frequency (the forward scan). The red curve is the spectrum obtained by sweeping the probe laser from high frequency to low frequency (the backward scan). The narrow peak on the left and the broad peak on the right correspond to transitions H2 and V2, respectively. We focus on the optical response from transition V2.

Ideally, the overall lineshape of transition V2 should be Lorentzian-like with a dark-state dip, as shown in Fig. 4a. However, the spectrum clearly shows a broadened lineshape with a round top and sharp edges, which is far from Lorentzian. The width and the strength of the dip corresponding to the dark state are also narrower and shallower than expected. More remarkably, we observe hysteresis at the sharp edges of the V2 absorption peak between the forward and backward scans. Additionally, the spectral position of the dark state in each scan is shifted in the same direction as the scan, which indicates a change of TPR when the scan direction is switched. The external magnetic field is unchanged in the forward and backward scans, so these observations indicate that we optically create and probe the dynamic nuclear spin polarization (DNP) in this charged quantum dot system, where the nuclear spin configuration depends on the laser sweeping direction.

The dependence on laser scan rate is shown in Fig. 1c. The dark state becomes more pronounced, concomitant with a broader dip width as we increase the laser scan rate (that is, the probe laser frequency is held for a shorter interval at each value). Under faster scans, the observed lineshape is closer to the standard dark-state spectrum in a Lambda level scheme (compare Fig. 4a). As we show below, the anomalous spectral features and their scan-rate dependence reflect the dynamical control of the nuclear field by the laser frequency scans on a timescale comparable to the nuclear spin relaxation time, which is of the order of a second[12,14,16].

We performed a set of measurements by fixing the frequencies of both lasers and recording the optical response as a function of time; these measurements reveal that the DNP modifies the Zeeman splitting, via the nuclear field, to maximize the trion excitation. Figure 2a shows the probe absorption spectra with forward (black) and backward (red) scans. We begin by scanning the laser backward and stopping the laser just before the sharp rising edge of the trion peak, as shown by the green curve in Fig. 2b. We record the absorption signal as a function of time with the laser frequency fixed. As shown in Fig. 2c, the system remains in hysteresis state 1 for a while (shown by the signal level) and then abruptly switches into hysteresis state 2, where it remains. This signifies that the nuclear field switches to a stable value that maximizes the trion excitation. Finally, we scan the probe laser forward and find that the subsequent partial forward scan spectrum (the blue curve in Fig. 2d) overlaps considerably with its equivalent in the full forward scan.

We also examined the dynamics of the nuclear spin by monitoring the dark state. As shown in Fig. 2e, after a full forward scan to locate the dark-state position (black curve), we took a partial forward scan

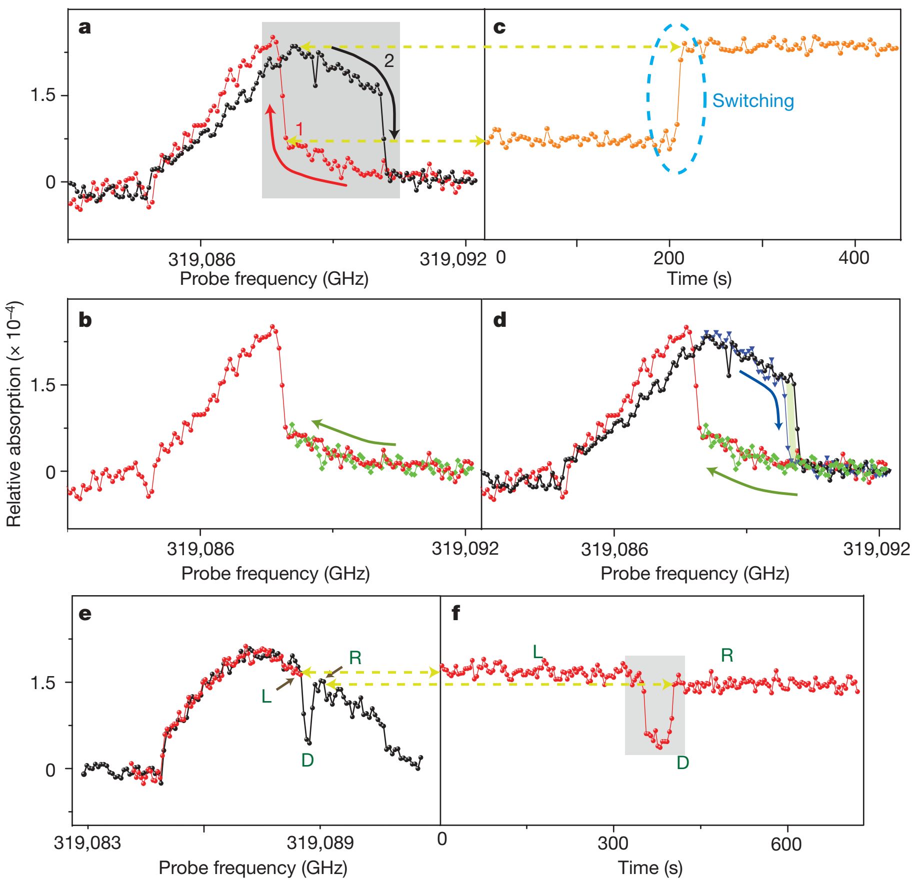


Figure 2 | Time-dependent probe absorption spectrum with fixed laser frequencies. Data are taken at a magnetic field of 2.64 T. **a**, The black (or red) curve represents a full forward (or backward) scan. **b**, The green curve is a partial backward scan. **c**, The probe absorption signal taken as a function of time immediately after stopping the laser just before the rising edge of the trion absorption. **d**, The blue curve is the partial forward scan taken after the switching of the hysteresis states. **e**, The black (or red) curve is the full forward (or partial forward) scan. L, D and R denote three system configurations. **f**, The absorption signal as a function of time, taken immediately after parking the laser just before the dark state is formed.

to prepare the initial nuclear spin configuration and stopped tuning the laser just before the formation of the dark state (red curve). Immediately, we measured the absorption signal as a function of time (Fig. 2f). The system starts in configuration L, indicated by the signal level, jumps into configuration D after some time, and then switches to configuration R, where it remains at high probe absorption. In experiments we noticed that the system can stay in the dark state D on a timescale from a few seconds to 3 min, indicating the meta-stable nature of the nuclear configuration at the TPR (see also the Supplementary Information). Figure 2f shows an example in which the system stays in the D configuration for ~40 s.

Power-dependent measurements of the dark state with a fast laser scan rate provide frequency domain information on the electron spin coherence time. Figure 3a shows the decreasing trend of the estimated spin decoherence rate $\gamma_s/2\pi$ with the increase of the square root of the pump intensities. The black dots represent $\gamma_s/2\pi$ inferred from the absorption minimum at the dark-state dip (normalized by the absorption maximum at the trion peak) and the blue triangles are values extracted from the best fit of the dark-state spectrum including the DNP dynamics. An example of the dark-state spectrum is given in Fig. 3b (identical to Supplementary Fig. S1 with pump Rabi of 0.9 GHz). Details of the data analysis leading to Fig. 3a can be found in the first section of Supplementary Information. For this time-ensemble-averaged measurement of a single spin, γ_s^{-1} corresponds to the inhomogeneous dephasing time T_2* because of the measurement-to-measurement fluctuations of the nuclear field. For our dot, we estimate the spin inhomogeneous broadening due to a thermally distributed nuclear environment to be (360 ± 30) MHz (refs 7, 21). The dark-state spectrum clearly shows that the spin T_2* has been enhanced well above the thermal value. As we show later, the enhancement of electron spin T_2* is a natural result from the suppression of nuclear spin fluctuations by the positive DNP feedback process and the data in Fig. 3a is a lower bound of this enhancement effect. This mechanism is in fact far more powerful in enhancing the spin T_2* than the preliminary results from Fig. 3a and b suggest, as we will show using a different experimental set-up.

Both the rounded and broadened trion peak in the probe spectrum and the switching behaviours at fixed laser frequencies indicate that large trion excitation is favoured by the DNP process. When a trion is excited, its two constituent electron spins form an inert singlet, leaving its hole constituent to interact with the nuclear spins. This is a unique element for optical control of the nuclear spin environment while manipulating the electron spin, and it accounts for the experimental observations here.

The hyperfine interaction between the hole spin and nuclear spin is strongly anisotropic[22,23]. In particular, it has a non-collinear hyperfine coupling term $S_h^X I_k^Z$, where S_h^X is the heavy-hole pseudospin operator along the field direction (**X**) and I_k^Z is the nuclear spin operator along the growth direction (**Z**). This interaction can flip a nuclear spin without flipping the hole spin, costing only the nuclear Zeeman energy: $\hbar\omega_N \approx 0.01\,\mathrm{GHz}\,\mathrm{T}^{-1}$ (refs 24 and 25). This process stands out from the various DNP interactions because the small energy cost can be directly compensated by the homogeneous broadening of the trion state (~0.4 GHz), which is a lower-order process than the electron-nuclear flipflops assisted by phonons or photons.

From Fermi's Golden Rule, the nuclear spin flip rates are proportional to $\rho_{t,i}\rho_{t,f}$, where $\rho_{t,i}$ is the initial-state trion population and $\rho_{t,f}$ is the final-state trion population after a nuclear spin flip (see details in Supplementary Information). This is because $S_h^X I_k^Z$ has non-zero matrix elements only in the trion portion of the steady-state wave function. The up and down nuclear spin flip rates are different, because they change the electron Zeeman splitting—and hence the two-photon detuning δ from the TPR—in opposite ways, each of which lead to different final-state trion populations $\rho_{t,f}$. Clearly, the

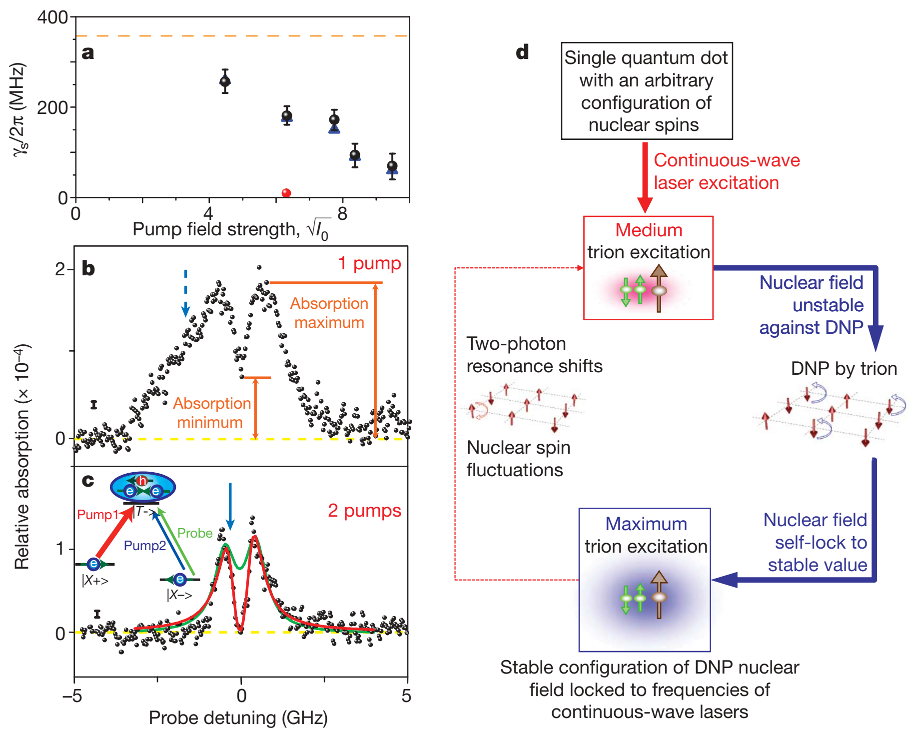


Figure 3 | The observation of the enhancement of electron spin T_2*. a, The estimated $\gamma_s/2\pi$ from the two-beam dark-state spectrum with various pump field strengths, where $I_o = 2\,\mathrm{W\,cm^{-2}}$. Black dots and blue triangles are estimated from the dip-to-peak absorption ratio with error bars determined from the measurement noise (see black error bar symbol in **b**) and from the best fits of the spectrum by modelling the nuclear field dynamics (see Supplementary Information), respectively. The dashed horizontal line denotes the thermal value. The red dot is from fitting the three-beam spectrum in **c**. **b**, An example of the dark-state spectra obtained by fast forward scan in the two-beam set-up that yields the $\gamma_s/2\pi$ value in **a**. The yellow dashed line denotes the zero signal line. **c**, A three-beam measurement with the schematic set-up shown in the inset. The solid blue arrow indicates the spectral position of the second pump beam (also indicated in **b** as a reference with the dashed blue arrow). **d**, Schematic illustration of the self-locking DNP feedback process, which locks the nuclear field to a stable value that maximizes trion excitation.

one resulting in a larger $\rho_{t,f}$ always 'wins', that is, the DNP process tends to maximize trion excitation. The net DNP rate is proportional to $\rho_t(\partial\rho_t/\partial\delta)$, which goes to zero at the maxima of the trion excitation (also the position of strongest absorption), located at $\delta=\pm\Omega_{\mathrm{pump}}/2$, where Ω_{pump} is the pump Rabi frequency (see Fig. 4b). The DNP process functions as a restoring force around the absorption maxima, such that when the laser is slightly detuned from the position of the strongest trion excitation, DNP acts to adjust the Zeeman splitting to maximize the trion excitation. We note that the net DNP rate is also zero at the TPR because it is a local minimum, which ties in with the meta-stable nature of the dark state observed in Fig. 2f.

We numerically simulate the self-locking process by including DNP into the optical Bloch equations (see Supplementary Information). Figure 4c simulates the result for fast scans, and Fig. 4d for slow scans, at magnetic field 1.32 T. The numerical simulations qualitatively reproduce the important features of the experimental observations. Figure 4e clearly shows that the resulting nuclear fields differ greatly depending on whether the probe frequency is scanned forward or backward, which explains the origins of the edge hysteresis and the spectral shift of the dark state. We note that the asymmetry between the forward and backward scans in Fig. 2a is due to the pump detuning (see Supplementary Fig. S3a).

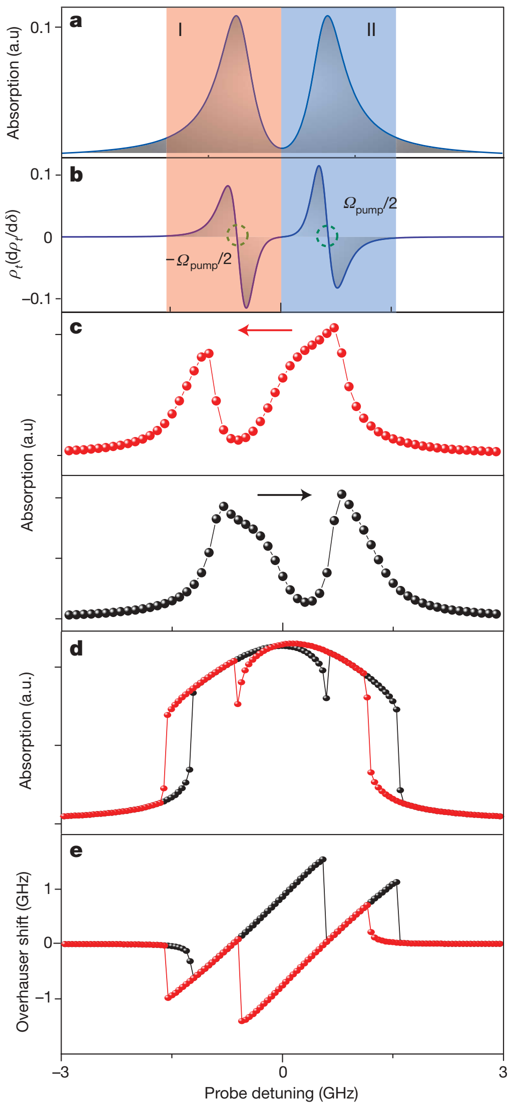


Figure 4 | Theoretical explanation of the nuclear field self-locking effect through the DNP feedback process. a, **b**, The calculated probe absorption spectrum (**a**) and $\rho_t(\partial\rho_t/\partial\delta)$ (**b**) by solving the three-level Lambda system with pump on resonance in the absence of DNP. **b** shows the DNP acting as a restoring force. The two stability regions corresponding to the absorption maxima are labelled I and II. **c**, **d**, Numerical simulation results including the self-locking DNP effects for fast (**c**) and slow (**d**) scan corresponding to a magnetic field of 1.32 T. The black (or red) curve represents the probe absorption spectrum of the forward (or backward) scan. **e**, The calculated nuclear field corresponding to the slow scan. The positive (or negative) nuclear field shifts the probe absorption spectrum to the blue (or red). a.u., arbitrary units.

The self-locking effect described in the theory also leads to the suppression of the nuclear spin fluctuations. Once the system has switched to a configuration of maximum trion excitation, the electron spin Zeeman energy and hence the nuclear field are determined and controlled only by the instantaneous laser frequencies, regardless of the initial nuclear spin configuration before the scan starts. In this regime, DNP can actively work to maximize the trion population, and any nuclear spin fluctuations that shift the Zeeman resonance are cancelled out through feedback via the DNP mechanism (Fig. 3d). The quantitative enhancement of T_2* by this mechanism is determined by the slope of the DNP rate as a function of detuning at the locking points, that is, the two circled positions in Fig. 4b at $\pm\Omega_{\mathrm{pump}}/2$. A larger slope means a stronger restoring force, and hence a better locking effect (see Supplementary Fig. S3b, which is qualitatively consistent with the trend shown in Fig. 3a). For spectra discussed in Fig. 3a and b, the locking position of the nuclear field follows the probe laser, which scans much faster than the DNP equilibration rate, so the suppression effect we obtained there is a lower bound of the capability of this nuclear field-locking technique.

Consequently, if the pump and probe beams are fixed spectrally to maximize the trion excitation, the nuclear field fluctuations should be suppressed further than when the probe continuously scans, as in Fig. 3b. Data taken using three beams (Fig. 3c) support this argument. The stronger pump 1 remains near-resonant with transition H1 and the weaker pump 2 is tuned to transition V2 and fixed at the spectral position that maximizes the trion absorption. The two pumps lock the nuclear field to a constant value and suppress nuclear fluctuations for the duration of the experiment. We use the weak probe beam to generate the dark-state spectrum with a fast scan rate, shown in Fig. 3c. The probe is weak and scans at a fast rate, so the effect of the probe beam on the nuclear field can be ignored.

The resulting spectrum in Fig. 3c shows a cleaner dark-state lineshape with a more pronounced dip than the two-beam dark-state spectrum in Fig. 3b with comparable pump intensity, which confirms that the nuclear field is locked by the two pumps. The dip strength represents the electron spin T_2* and the measured absorption at the TPR approaches zero, so the data indicate a substantial enhancement of the T_2* in this two-pump set-up. It is challenging to extract an accurate spin decoherence rate because the suppression of nuclear spin fluctuations is so strong that the signal approaches zero at the TPR. However, fitting the data with the standard two-beam optical Bloch equation, the red curve on top of the data (fitting parameters $\Omega_{\mathrm{pump}}/2\pi=0.9$ GHz, $\gamma_t/2\pi=0.4$ GHz and $\delta_{\mathrm{pump}}=-30$ MHz), yields a value for $\gamma_s/2\pi$ of the order of ~1 MHz with a 5 MHz upper bound error bar, limited by the mutual coherence bandwidth between the lasers. We can also estimate the T_2* directly from the absorption minimum at the dark-state dip, although this will not be as accurate as the optical Bloch equation curve fitting because it does not exploit all the data points. This dip-to-peak ratio estimation gives a $\gamma_s/2\pi$ of 2 MHz, which agrees with the optical Bloch equation fit. The green curve on top of the data is a theoretical plot using the thermal value of T_2*, which clearly shows the strong enhancement of the electron spin T_2* by the DNP self-locking effect. The strongly suppressed γ_s indicates that the intrinsic T_2 could possibly be recovered by this nuclear field-locking technique. We expect that a time domain measurement may more accurately show the dramatic enhancement of T_2* by the nuclear field-locking scheme described above. Further measurements and theory to explore this potential are in progress.

In conclusion, our results provide a simple but powerful method of suppressing the nuclear spin fluctuations. Once the nuclear spin environment is prepared by our nuclear field-locking method, it may be possible to perform coherent manipulations of a single electron spin for the entire duration of the intrinsic electron spin coherence time, unaffected by hyperfine-interaction-induced inhomogeneous spin dephasing.

METHODS SUMMARY

We give a brief review of the sample structure and experimental techniques, which are explained in detail in refs 26 and 27. The sample contains InAs self-assembled quantum dots grown by molecular beam epitaxy. The size of the dot is about 3 nm in height and 15 nm in base diameter[28]. The sample is embedded in a Schottky diode structure, so we set the DC voltage to charge only one electron into the dot. We use the DC Stark shift modulation absorption technique to achieve a high signal-to-noise ratio[29]. The modulation amplitude is large so that the data correspond directly to absorption. The sample is held in a continuous helium flow magneto cryostat. The magnetic field can be tuned up to 6.6 T. For the quantum dot of interest, the electron in-plane g factor is 0.49 and the hole in-plane g factor is 0.13. The experiment is performed at a temperature of $\sim$5 K.

Two continuous-wave lasers are used in the experiment. As shown in Fig. 1a, a strong pump (red arrow) is horizontally polarized and fixed to be nearly resonant with transition H1 and the weak probe beam (green arrow) is vertically polarized. In the experiment, the estimated pump Rabi frequency ranges from $\sim$0.63 to $\sim$1.35 GHz, and the probe Rabi frequency is fixed at $\sim$0.24 GHz. Since the polarization axis of the quantum dot is rotated about 20° away from the laboratory frame owing to the heavy and light hole mixing[30], the probe beam can pick up both V2 and H2 transitions in a single scan. For the slow scan data shown in the text, the probe laser frequency is held at each value for 4 s and each data point is the integrated signal over the last 1 s of this interval.

Received 30 November 2008; accepted 5 May 2009.

1. Berezovsky, J. *et al.* Picosecond coherent optical manipulation of a single electron spin in a quantum dot. *Science* **320,** 349–352 (2008).
2. Bracker, A. S. *et al.* Optical pumping of the electronic and nuclear spin of single charge-tunable quantum dots. *Phys. Rev. Lett.* **94,** 047402 (2005).
3. Gammon, D. & Steel, D. G. Optical studies of single quantum dots. *Phys. Today* **55,** 36–41 (2002).
4. Gerardot, B. D. *et al.* Optical pumping of a single hole spin in a quantum dot. *Nature* **451,** 441–444 (2008).
5. Kim, D. *et al.* and nondestructive measurement in a quantum dot molecule. *Phys. Rev. Lett.* **101,** 236804 (2008).
6. Kroutvar, M. *et al.* Optically programmable electron spin memory using semiconductor quantum dots. *Nature* **432,** 81–84 (2004).
7. Merkulov, I. A. Efros, A. L. & Rosen, M. Electron spin relaxation by nuclei in semiconductor quantum dots. *Phys. Rev. B* **65,** 205309 (2002).
8. Petta, J. R. *et al.* Coherent manipulation of coupled electron spins in semiconductor quantum dots. *Science* **309,** 2180–2184 (2005).
9. Baugh, J. Kitamura, Y. Ono, K. & Tarucha, S. Large nuclear Overhauser fields detected in vertically coupled double quantum dots. *Phys. Rev. Lett.* **99,** 096804 (2007).
10. Tartakovskii, A. I. *et al.* Nuclear spin switch in semiconductor quantum dots. *Phys. Rev. Lett.* **98,** 026806 (2007).
11. Eble, B. *et al.* Dynamic nuclear polarization of a single charge-tunable InAs/GaAs quantum dot. *Phys. Rev. B* **74,** 081306 (2006).
12. Greilich, A. *et al.* Nuclei-induced frequency focusing of electron spin coherence. *Science* **317,** 1896–1899 (2007).
13. Koppens, F. H. L. *et al.* Control and detection of singlet-triplet mixing in a random nuclear field. *Science* **309,** 1346–1350 (2005).
14. Maletinsky, P., Lai, C. W., Badolato, A. & Imamoglu, A. Nonlinear dynamics of quantum dot nuclear spins. *Phys. Rev. B* **75,** 035409 (2007).
15. Vink, I. T. *et al.* Locking electron spins into magnetic resonance by electron-nuclear feedback. Preprint at <http://arxiv.org/abs/0902.2659> (2009).
16. Reilly, D. J. *et al.* Suppressing spin qubit dephasing by nuclear state preparation. *Science* **321,** 817–821 (2008).
17. Xu, X. D. *et al.* Fast spin state initialization in a singly charged InAs-GaAs quantum dot by optical cooling. *Phys. Rev. Lett.* **99,** 097401 (2007).
18. Xu, X. *et al.* Coherent population trapping of an electron spin in a single negatively charged quantum dot. *Nature Phys.* **4,** 692–695 (2008).
19. Harris, S. E. Electromagnetically induced transparency. *Phys. Today* **50,** 36–42 (1997).
20. Stepanenko, D., Burkard, G., Giedke, G. & Imamoglu, A. Enhancement of electron spin coherence by optical preparation of nuclear spins. *Phys. Rev. Lett.* **96,** 136401–136404 (2006).
21. Braun, P. F. *et al.* Direct observation of the electron spin relaxation induced by nuclei in quantum dots. *Phys. Rev. Lett.* **94,** 116601 (2005).
22. Eble, B. *et al.* Hole–nuclear spin interaction in quantum dots. *Phys. Rev. Lett.* **102,** 146601–146604 (2009).
23. Fischer, J. Coish, W. A., Bulaev, D. V. & Loss, D. Spin decoherence of a heavy hole coupled to nuclear spins in a quantum dot. *Phys. Rev. B* **78,** 155329 (2008).
24. Kikkawa, J. M. & Awschalom, D. D. All-optical magnetic resonance in semiconductors. *Science* **287,** 473–476 (2000).
25. Gammon, D. *et al.* Nuclear spectroscopy in single quantum dots: nanoscopic raman scattering and nuclear magnetic resonance. *Science* **277,** 85–88 (1997).
26. Ware, M. E. *et al.* Polarized fine structure in the photoluminescence excitation spectrum of a negatively charged quantum dot. *Phys. Rev. Lett.* **95,** 177403 (2005).
27. Xu, X. *et al.* Coherent optical spectroscopy of a strongly driven quantum dot. *Science* **317,** 929–932 (2007).
28. Scheibner, M. *et al.* Optically mapping the electronic structure of coupled quantum dots. *Nature Phys.* **4,** 291–295 (2008).
29. Alen, B. *et al.* Stark-shift modulation absorption spectroscopy of single quantum dots. *Appl. Phys. Lett.* **83,** 2235–2237 (2003).
30. Koudinov, A. V. & Akimov, I. A. Kusrayev, Yu. G. & Henneberger, F. Optical and magnetic anisotropies of the hole states in Stranski-Krastanov quantum dots. *Phys. Rev. B* **70,** 241305 (2004).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank P. L. McEuen, L.-M. Duan, and D. Kim for discussions. This work is supported by US ARO, AFOSR, ONR, NSA/LPS, and FOCUS-NSF.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to D.G.S. (dst@umich.edu).

Vol 459|25 June 2009|doi:10.1038/nature08115

LETTERS

High-speed tracking of rupture and clustering in freely falling granular streams

John R. Royer[1], Daniel J. Evans[1], Loreto Oyarte[1], Qiti Guo[1], Eliot Kapit[1], Matthias E. Möbius[1]†, Scott R. Waitukaitis[1] & Heinrich M. Jaeger[1]

Thin streams of liquid commonly break up into characteristic droplet patterns owing to the surface-tension-driven Plateau–Rayleigh instability[1–3]. Very similar patterns are observed when initially uniform streams of dry granular material break up into clusters of grains[4–6], even though flows of macroscopic particles are considered to lack surface tension[7,8]. Recent studies on freely falling granular streams tracked fluctuations in the stream profile[9], but the clustering mechanism remained unresolved because the full evolution of the instability could not be observed. Here we demonstrate that the cluster formation is driven by minute, nanoNewton cohesive forces that arise from a combination of van der Waals interactions and capillary bridges between nanometre-scale surface asperities. Our experiments involve high-speed video imaging of the granular stream in the co-moving frame, control over the properties of the grain surfaces and the use of atomic force microscopy to measure grain–grain interactions. The cohesive forces that we measure correspond to an equivalent surface tension five orders of magnitude below that of ordinary liquids. We find that the shapes of these weakly cohesive, non-thermal clusters of macroscopic particles closely resemble droplets resulting from thermally induced rupture of liquid nanojets[10–12].

Granular systems in Hele–Shaw geometries or impinging on stationary targets have recently been shown to provide excellent approximations of liquid behaviour in the limit approaching zero surface tension[7,8]. In these experiments, however, the material was strongly forced and momentum transfer dominated other forces. To observe the vestiges of any residual surface tension and estimate its magnitude requires conditions where the material experiences as little external forcing as possible. Motivated by the strikingly liquid-like appearance of streams of fine-grained granular material breaking into droplets, first reported in 1890 (ref. 13) and more recently observed within the context of granular jets[4,5], we investigated freely falling granular streams, analogous to liquids falling from a faucet. These streams provide a controlled system for the study of droplet formation provided that break-up events can be tracked with high spatial resolution at high imaging speeds. The difficulty is that the time to break-up increases rapidly with decreasing surface tension. For granular streams this requires tracking grains for half a metre or more before they form clusters. We achieve this using a high-speed video camera that falls with the stream and report here on experiments capturing the granular break-up dynamics in detail.

Figure 1 shows the break-up process for a stream of glass spheres emerging from a small opening in the bottom of a hopper mounted inside a 2.5-m-tall, evacuated cylinder (Fig. 1a). As the stream accelerates under gravity, an axial velocity gradient develops and elongates the stream as it falls. While the stream stretches, initial

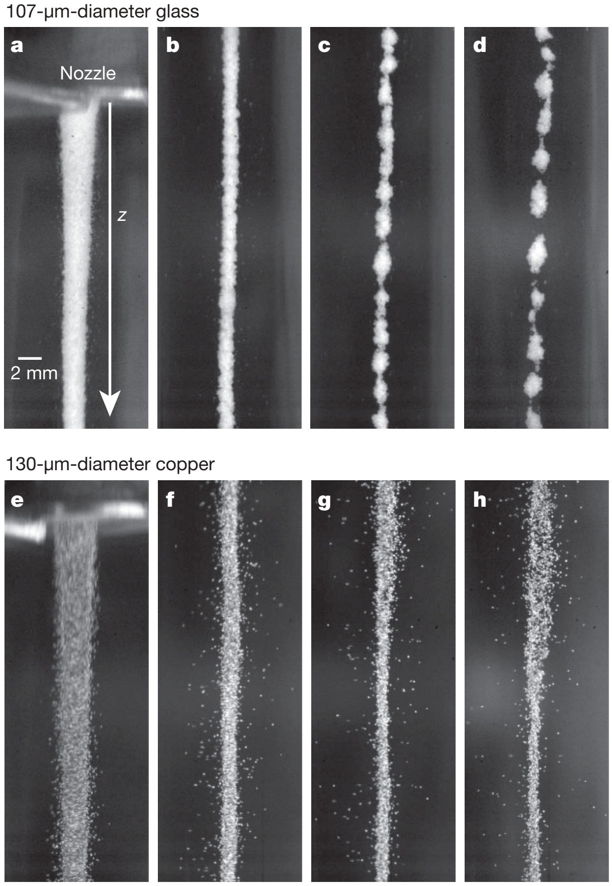


Figure 1 | Break-up of a granular stream. a–d, Stream of glass grains of diameter $d = 107\ \mu m \pm 19\ \mu m$ falling out of a nozzle of diameter $D_0 = 4.0$ mm at the following points: just below the nozzle (**a**), and $z = 20$ cm (**b**), $z = 55$ cm (**c**) and $z = 97$ cm (**d**) from the top of the frame to the nozzle. **e–h**, Stream of copper grains of diameter $d = 130\ \mu m \pm 30\ \mu m$ falling out of the same $D_0 = 4$ mm nozzle at the following points: just below the nozzle (**e**), and $z = 20$ cm (**f**), $z = 55$ cm (**g**) and $z = 97$ cm (**h**) from the top of the frame to the nozzle. The nozzle and reservoir of grains are housed in a 2.5-m-tall acrylic tube, which is sealed and evacuated to 0.03 kPa (gas mean free path, ~200 μm) to reduce air drag. We use a high-speed camera (Phantom v7.1) falling along a low-friction rail outside of the chamber to track a 3-cm-long section of the stream as it falls from the nozzle to the bottom of the chamber (0.04 mm per pixel, 1,000 frames per second; see Supplementary Movies). An optical encoder measures the position of the camera, allowing us to correct for small deviations from free-fall as the camera moves along the rail.

[1]James Franck Institute and Department of Physics, The University of Chicago, Chicago, Illinois 60637, USA. †Present address: School of Physics, Trinity College Dublin, Dublin 2, Ireland.

undulations emerge and deepen (Fig. 1b), creating clusters connected by thin bridges a few grains wide (Fig. 1c). These bridges eventually rupture as the clusters continue to separate (Fig. 1d). This ability to track the evolution of individual clusters over several metres and image the rupture details provides unique insights into the underlying clustering mechanism.

Whereas the instability of ordinary liquid columns is driven by molecular surface tension, possible mechanisms for droplet formation in granular systems include hydrodynamic interactions with the surrounding gas, inelastic grain–grain collisions, and cohesive forces. Hydrodynamic interactions have indeed recently been associated with fluctuations in the profile of streams falling in air[9]; however, from experiments across a wide range of ambient pressures down to 0.03 kPa we find that grain–gas interactions do not drive clustering (Supplementary Fig. S1), in agreement with earlier work[6]. Inelastic collisions are known to produce clusters in other granular systems[14–16], although for realistic grains, where the coefficient of restitution approaches unity at low velocities, these clusters may be transient. To vary the inelasticity, we replaced the glass spheres by copper grains of similar size. The grain mass does not enter scenarios solely driven by inelasticity, so the smaller coefficient of restitution (for copper[17], $e \approx 0.90$, for soda lime glass[18], $e \approx 0.97$ at impact velocities of 0.5–1.0 m s^{-1}) should lead to even more pronounced clustering. As shown in Fig. 1e–h, the opposite behaviour is observed: Instead of breaking up into discrete, compact clusters, sections of the stream begin to drift apart and expand in the radial direction.

In principle, cohesion might arise from a variety of sources, including electrostatic charging, capillary or van der Waals forces[19–21]. For a rough estimate of the cohesive strength we track clusters as they fall and accelerate to a speed at which Stokes drag pulls individual grains off cluster protrusions. Correcting for slight changes in the air viscosity at reduced pressure, this gives values of a few nanoNewtons. To compare this to any electrostatic forces present, we obtain the distribution of charges on the grains by applying a uniform electric field perpendicular to the falling stream and tracking individual grain trajectories (see Supplementary Information). For both glass and copper, we find the streams are neutral overall but contain a small fraction of positively and negatively charged grains, up to a roughly $q_{max} = \pm 100{,}000$ electron charges per grain (Supplementary Fig. S2). Still, this gives attractive electrostatic forces a maximum $F_{max} = (1/4\pi\varepsilon_0)q_{max}^2/d^2 \approx 0.1$ nN for grains with diameter $d = 100\,\mu$m, too weak to be the dominant cohesive force. (Here $\varepsilon_0 = 8.85 \times 10^{-12}$ C^2 N^{-1} m^{-2} is the permittivity of free space.) Furthermore, experiments with conductive, silver-coated 100-μm-diameter glass spheres produce clusters identical to experiments using uncoated spheres, emphasizing that electrostatic forces do not drive the observed clustering.

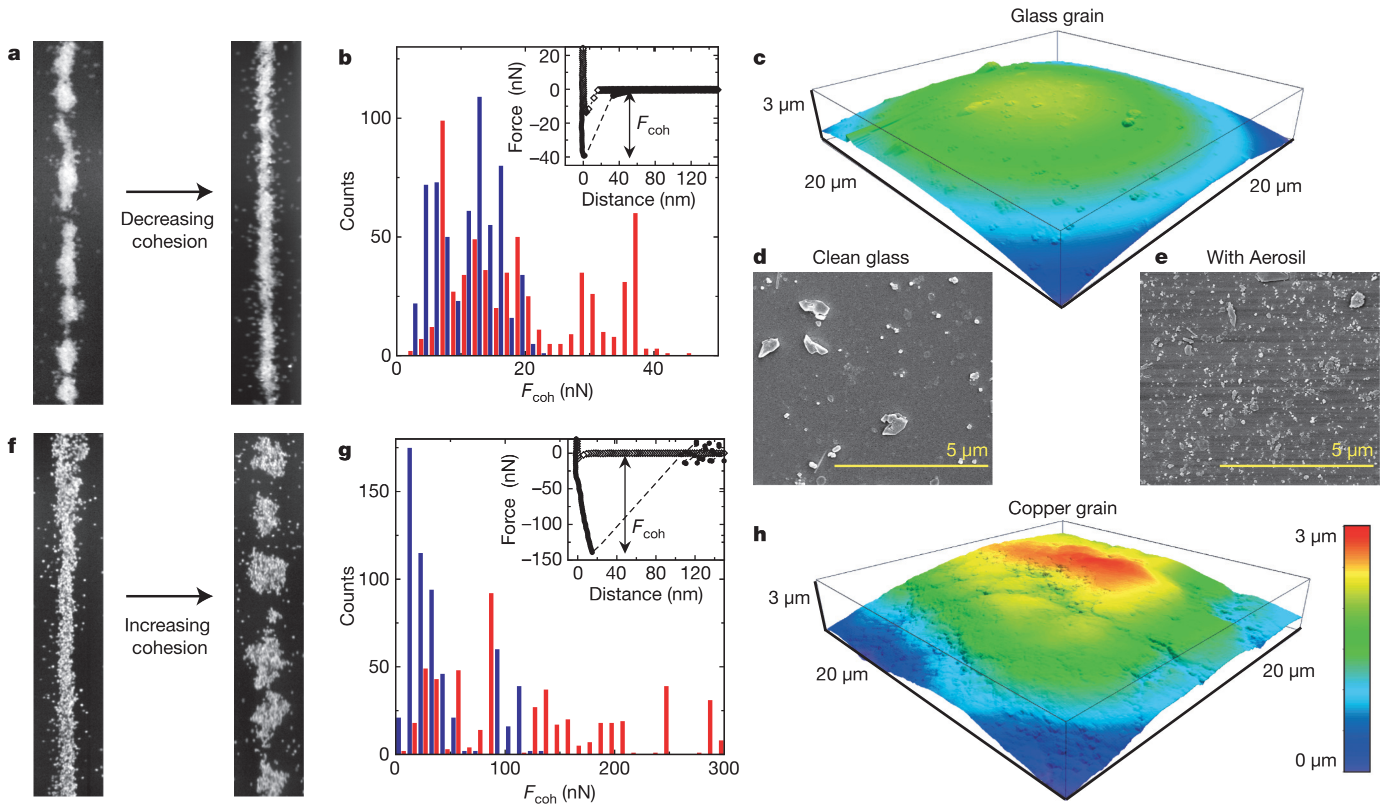


Figure 2 | Controlling clustering by altering cohesion. **a**, Stream of glass grains of diameter $d = 151\,\mu\text{m} \pm 26\,\mu\text{m}$ (left) and the same grains coated with silica particles of size 10–100 nm (Aerosil 200, Degussa) (right). Aerosil is a common additive used to improve the flowability of powders by reducing cohesion. The addition of Aerosil altered but did not eliminate clustering in finer glass grains of diameter 107 μm or 54 μm. **b**, Histograms of cohesive forces between pairs of grains for untreated, clean glass grains (blue) and Aerosil-coated grains (red). Forces were measured using an Asylum Research MFP-3D-Bio AFM. Single grains were epoxy-glued to the AFM cantilever (Nanosensors, spring constant $k \approx 2$ N m^{-1}) and brought into contact with fixed grains of the same material. The inset to **b** shows a typical force–displacement curve for a pair of untreated glass grains, showing the approach (open symbols) and retraction (closed symbols) of the cantilever. F_{coh} is the maximum force magnitude before two grains abruptly snap apart during retraction (dotted lines). Each histogram consists of measurements from 20 separate force curves from 15 different grains. **c**, AFM topographic map of 20 μm × 20 μm section of an untreated glass grain. **d**, **e**, Scanning electron microscope images of an untreated glass grain (**d**) and a grain coated with Aerosil (**e**). The grains were coated by mixing a small amount (~0.1% by weight) of Aerosil with untreated grains. Scanning electron microscope imaging of random samples confirmed that the Aerosil was well dispersed across the grain surfaces. **f**, Stream of copper grains of diameter $d = 130\,\mu\text{m} \pm 30\,\mu\text{m}$ (left) and the same copper mixed with a small amount of oil (right). **g**, Histogram of pull-off forces for untreated grains (blue) and oil coated grains (red). The inset to **g** shows a force–displacement curve illustrating the larger pull-off force from oil-coated grains and the measurable displacement of the grain before it snaps off the surface. **h**, AFM topographic map of a 20 μm × 20 μm section of a copper grain.

To determine the intergrain cohesive forces F_{coh} directly, we record force–displacement curves of individual grains brought into contact and pulled apart by an atomic force microscope (AFM)[22,23] (Fig. 2). F_{coh} varies across different grains, but is generally a few tens of nanoNewtons, in agreement with our earlier estimate. Both van der Waals and capillary forces for perfectly smooth spheres in contact scale as $F_{coh} \propto \gamma d$, with interfacial energy γ of the order of tens of milliNewtons per metre (ref. 19), giving microNewton forces for grains of diameter $d = 150\,\mu m$. The orders-of-magnitude-smaller forces observed (Fig. 2b, g), as well as the absence of any scaling of F_{coh} with grain diameter, emphasize the significance of the local, nanometre-scale grain topography[23,24]. It is difficult to distinguish van der Waals from capillary forces because we cannot rule out molecularly thin absorbed films that create tiny bridges between individual asperities[24,25]. However, we still observe clustering in glass grains stored under vacuum (0.05 kPa) at low humidity (<1%) and also in grains coated with hydrophobic silane.

To confirm that tiny, short-ranged cohesive forces are responsible for the clustering, we can decrease their strength by enhancing the grains' surface roughness. As shown in Fig. 2 for 150-µm-diameter glass grains, adding nanoscale asperities to the grain surfaces reduces F_{coh} by roughly a factor of two and indeed reduces clustering significantly. However, what controls clustering cannot be determined solely from knowledge of F_{coh}. In fact, although clean glass does form clusters, F_{coh} for glass is not larger than for copper grains with similar mass (median values are 17 nN and 30 nN, respectively; see Fig. 2). One reason is that AFM pull-off force measurements only mimic head-on collisions, while both normal and tangential force components must be considered to describe fully the collision of macroscopic bodies. Topographic maps show significant differences in the nature of the roughness (Fig. 2c, h), and the much more pronounced large-scale roughness for copper is likely to reduce sliding, leading to different rotational collision dynamics. For coarse sand, rough and irregular like copper, we indeed observe weaker, less compact clusters (Supplementary Information).

We can greatly simplify the picture and eliminate many of the complications associated with dissipation due to sliding or rolling when 'sticky' collisions are the dominant source of energy loss. In this scenario, inelastic interactions initially reduce the relative particle velocities δu to a level where colliding grains with mass m get captured

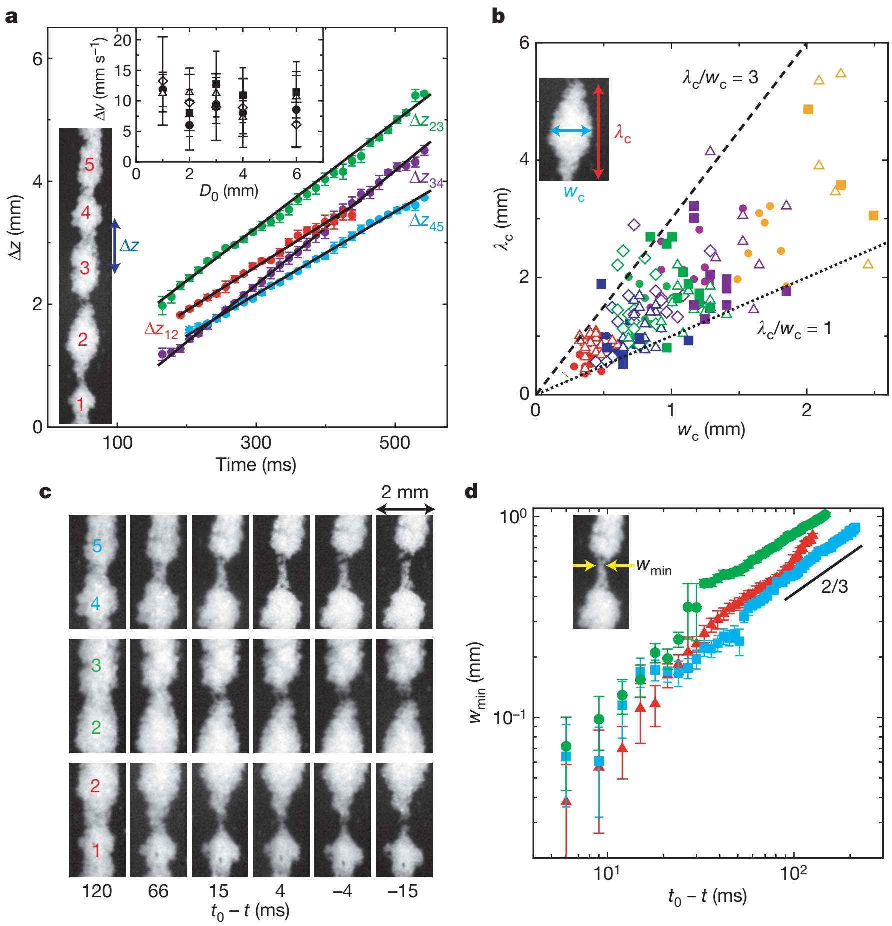


Figure 3 | Clustering dynamics. **a**, Separation of clusters Δz versus time t (see diagram) for representative sample of five adjacent clusters from the stream shown in Fig. 1a–d (107-µm-diameter glass, $D_0 = 4$ mm). Δz was measured by tracking local maxima in the width of the stream both before and after the clusters separated. $t = 0$ corresponds to the time where each section of the stream left the nozzle. Solid lines show fits to $\Delta z(t) = \Delta z_0 + (\Delta v)t$. Linear growth was observed for all nozzle diameters and grain materials. The inset to **a** shows the median separation velocity Δv obtained from fits to $\Delta z(t)$ versus nozzle diameter D_0 for glass spheres of $d = 54\,\mu m \pm 10\,\mu m$ (open triangles), $d = 107\,\mu m \pm 19\,\mu m$ (solid circles), $d = 151\,\mu m \pm 26\,\mu m$ (open diamonds), and oil-coated copper particles of $d = 130\,\mu m \pm 30\,\mu m$ (solid squares). **b**, Cluster length λ_c versus cluster width w_c (see diagram) for individual clusters. Different grain materials are represented by the same symbols as in **a**. Different nozzle diameters are colour-coded: $D_0 = 1.0$ mm (red), 2.0 mm (blue), 2.9 mm (green), 4.0 mm (purple) and 6.1 mm (orange). **c**, Thinning and rupture of the neck between adjacent clusters for three of the pairs of clusters from **a**. Rupture time t_0 is the first frame where an open gap appears between clusters. From the time each section of the stream left the nozzle, $t_0 = 421$ ms (pair 4–5), 325 ms (pair 2–3) and 350 ms (pair 1–2). **d**, Log–log plot of the minimum neck diameter w_{min} (see diagram) versus time to rupture for each of the pairs of clusters in **c**: pair 4–5 (blue), pair 2–3 (green) and pair 1–2 (red). Solid line with a slope of +2/3 drawn for comparison. Error bars on Δz and w_{min} indicate statistical errors from averaging over multiple frames. Error bars on Δv show the standard deviation based on 6–10 pairs of clusters from a single experiment.

by cohesive trapping potentials. This happens when the pre-collision kinetic energy or 'granular temperature' $(m/2)(\delta u)^2$ becomes smaller than $W_{\mathrm{coh}} = \int F_{\mathrm{coh}}(\xi)\mathrm{d}\xi$, the work done against a cohesive force $F_{\mathrm{coh}}(\xi)$ as two particles try to separate themselves by a distance ξ (refs 26, 27). This energy loss to cohesion prevents the coefficient of restitution from approaching unity at low velocities, instead causing it to rapidly drop to zero as the collision velocity approaches $v_{\mathrm{c}} = (2W_{\mathrm{coh}}/m)^{1/2}$.

This scenario can be tested quantitatively by using rough grains (to reduce sliding) and adding cohesion (to induce sticking). To do this, we mixed copper grains with a small amount ($\sim 3 \times 10^{-4}$ by volume) of mineral oil, amounting to a 10-nm-thick coat on average. As Fig. 2 shows, this results in a fivefold increase of the median pull-off force for copper, which now clusters very similarly to glass. Tracing $F_{\mathrm{coh}}(\xi)$ over 10–20 nm we calculate $W_{\mathrm{coh}} \approx 10^{-15}$ J. The velocity threshold for sticking collisions between two copper grains (density $\rho = 8.9\,\mathrm{kg\,m^{-3}}$, $d = 130\,\mu\mathrm{m}$) is then $v_{\mathrm{c}} = (2W_{\mathrm{coh}}/m)^{1/2} \approx 0.4\,\mathrm{mm\,s^{-1}}$. This value agrees well with upper limits on the relative particle velocities inside clusters obtained from direct imaging. For example, for the glass data in Fig. 1, from the appearance of undulations (Fig. 1b) to the fully formed clusters (Fig. 1d) 240 ms later, the maxima in the stream width change by <0.1 mm, limiting δu to less than $0.5\,\mathrm{mm\,s^{-1}}$ inside clusters. It also explains the absence of clustering when the velocity fluctuations are much larger, for example, at the nozzle where $\delta u \approx 1\,\mathrm{cm\,s^{-1}}$ for both glass and copper as determined directly by particle imaging velocimetry.

We can use our results to estimate an equivalent 'granular' surface tension γ. In a conventional liquid, surface tension scales as $\gamma \approx \varepsilon/\sigma^2$, where ε is the depth of the attractive potential between molecules and σ is their size[28]. Using W_{coh} instead of ε and d instead of σ, we find $\gamma \approx 0.1\,\mu\mathrm{N\,m^{-1}}$, which is five orders of magnitude lower than in conventional liquids. To the extent that the granular streams behave like liquids, our experiments can therefore probe an ultralow-surface-tension regime that has recently attracted attention—such as simulations of nanojets[10–12], and colloidal systems a few molecules or particles wide[29]. For liquids in this regime, thermal fluctuations are sufficient to drive break-up. Intriguingly, the granular cluster shapes exhibit striking similarities, including characteristic double-cone necks at break-up.

To explore these connections quantitatively, we track details of the break-up between pairs of clusters (Fig. 3). The separation velocity between any two clusters varies significantly, as does the cluster size and the precise shape of the necks; on average, however, separation speed, cluster aspect ratio and neck evolution exhibit remarkably robust behaviour, largely independent of grain type, grain size or nozzle diameter. Independently of surface tension, in both macroscopic and nanoscale liquid streams, perturbations are unstable only for wavelengths λ greater than the stream circumference πw (refs 2, 10, 12). For all granular streams studied, the ratio of cluster length λ_{c} to cluster width w_{c} instead appears to be smaller, falling between one and three (Fig. 3b). This narrow interval is independent of experimental details, which might explain why similar clusters also form under quite different initial conditions, for example, in granular jets formed by sphere impact[4,5]. Finally, Fig. 3c and d shows the evolution of three different necks taken from a small section of the granular stream. In liquid nano-threads the minimum neck width is predicted[30] to scale as $w_{\mathrm{min}} \propto (t_0 - t)^{0.418}$ with time to break-off at t_0. While our data range is too small to claim any particular scaling, the data nevertheless appear incompatible with exponents $<1/2$. At early times the data seem much more in line with inviscid break-up (exponent 2/3) of macroscopic liquid streams[2], although there the neck shapes are quite different[1].

These experimental results open up new territory for which there is at present no theoretical framework. The granular streams act like dense, cold fluids. The low temperature allows us to observe the exceedingly weak cohesive forces responsible for clustering and extract an effective, ultralow surface tension. Although the break-up shapes are similar to those seen in simulations of mesoscopic liquid threads such as nanojets, the clustering results from collisional cooling rather than thermal fluctuations. Given that freely falling granular streams are exquisitely sensitive probes for minute forces they also provide a new tool with which to measure cohesive energies in granular systems.

Received 22 February; accepted 5 May 2009.

1. Shi, X. D., Brenner, M. P. & Nagel, S. R. A cascade of structure in a drop falling from a faucet. *Science* **265**, 219–222 (1994).
2. Eggers, J. & Villermaux, E. Physics of liquid jets. *Rep. Prog. Phys.* **71**, 036601 (2008).
3. Doshi, P. *et al.* Persistence of memory in drop breakup: the breakdown of universality. *Science* **302**, 1185–1188 (2003).
4. Lohse, D. *et al.* Impact on soft sand: void collapse and jet formation. *Phys. Rev. Lett.* **93**, 198003 (2004).
5. Royer, J. R. *et al.* Formation of granular jets observed by high-speed X-ray radiography. *Nature Phys.* **1**, 164–167 (2005).
6. Möbius, M. E. Clustering instability in a freely falling granular jet. *Phys. Rev. E* **74**, 051304 (2006).
7. Cheng, X., Xu, L., Patterson, A., Jaeger, H. M. & Nagel, S. R. Towards the zero-surface-tension limit in granular fingering instability. *Nature Phys.* **4**, 234–237 (2008).
8. Cheng, X., Varas, G., Citron, D., Jaeger, H. M. & Nagel, S. R. Collective behavior in a granular jet: emergence of a liquid with zero surface tension. *Phys. Rev. Lett.* **99**, 188001 (2007).
9. Amarouchene, Y., Boudet, J.-F. & Kellay, H. Capillarylike fluctuations at the interface of falling granular jets. *Phys. Rev. Lett.* **100**, 218001 (2008).
10. Koplik, J. & Banavar, J. R. Molecular dynamics of interface rupture. *Phys. Fluids A* **5**, 521–536 (1993).
11. Moseler, M. & Landman, U. Formation, stability, and breakup of nanojets. *Science* **289**, 1165–1169 (2000).
12. Kawano, S. Molecular dynamics of rupture phenomena in a liquid thread. *Phys. Rev. E* **58**, 4468–4472 (1998).
13. Khamontoff, N. Application of photography to the study of the structure of trickles of fluid and dry materials. *J. Russ. Phys-Chem. Soc.* **22**, 281–284 (1890).
14. Goldhirsch, I. Rapid granular flows. *Annu. Rev. Fluid Mech.* **35**, 267–293 (2003).
15. Brilliantov, N. V. & Pöschel, T. *Kinetic Theory of Granular Gases* (Oxford University Press, 2004).
16. Efrati, E., Livne, E. & Meerson, B. Hydrodynamic singularities and clustering in a freely cooling inelastic gas. *Phys. Rev. Lett.* **94**, 088001 (2005).
17. Kuwabara, G. & Kono, K. Restitution coefficient in a collision between two spheres. *Jpn. J. Appl. Phys.* **26**, 1230–1233 (1987).
18. Foerster, S. F., Louge, M. Y., Chang, H. & Allia, K. Measurements of the collision properties of small spheres. *Phys. Fluids* **6**, 1108–1115 (1994).
19. Israelachvili, J. N. *Intermolecular and Surface Forces* 2nd edn (Academic Press, 1992).
20. Podczeck, F. *Particle-Particle Adhesion in Pharmaceutical Powder Handling* (Imperial College Press, 1998).
21. Visser, J. Van der Waals and other cohesive forces affecting powder fluidization. *Powder Technol.* **58**, 1–10 (1989).
22. Jones, R. From single particle AFM studies of adhesion and friction to bulk flow: forging the links. *Granular Matter* **4**, 191–204 (2003).
23. Schaefer, D. M. *et al.* in *Fundamentals of Adhesion and Interfaces* (eds Rimai, D. S., Demejo, L. P. & Mittal, K. L.) 35–48 (VSP, 1995).
24. Halsey, T. C. & Levine, A. J. How sandcastles fall. *Phys. Rev. Lett.* **80**, 3141–3144 (1998).
25. Bocquet, L., Charlaix, E., Ciliberto, S. & Crassous, J. Moisture-induced ageing in granular media and the kinetics of capillary condensation. *Nature* **396**, 735–737 (1998).
26. Brilliantov, N. V., Albers, N., Spahn, F. & Pöschel, T. Collision dynamics of granular particles with adhesion. *Phys. Rev. E* **76**, 051302 (2007).
27. Sorace, C. M., Louge, M. Y., Crozier, M. D. & Law, V. H. C. High apparent adhesion energy in the breakdown of normal restitution for binary impacts of small spheres at low speed. *Mech. Res. Commun.* **36**, 364–368 (2009).
28. Rowlinson, J. S. & Widom, B. *Molecular Theory of Capillarity* (Clarendon Press, 1982).
29. Hennequin, Y. *et al.* Drop formation by thermal fluctuations at an ultralow surface tension. *Phys. Rev. Lett.* **97**, 244502 (2006).
30. Eggers, J. Dynamics of liquid nanojets. *Phys. Rev. Lett.* **89**, 084502 (2002).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank X. Cheng, R. Cocco, E. Corwin, R. Karri, N. Keim, T. Knowlton, S. Nagel, T. Witten and W. Zhang for discussions and J. Jureller for AFM training and assistance. This work was supported by NSF through its MRSEC programme and the Inter-American Materials Collaboration Chicago-Chile, and by the Keck Initiative for Ultrafast Imaging at the University of Chicago.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to J.R.R. (jroyer@uchicago.edu).

LETTERS

Enhanced carbon pump inferred from relaxation of nutrient limitation in the glacial ocean

L. E. Pichevin[1], B. C. Reynolds[2], R. S. Ganeshram[1], I. Cacho[3], L. Pena[3], K. Keefe[4] & R. M. Ellam[4]

The modern Eastern Equatorial Pacific (EEP) Ocean is a large oceanic source of carbon to the atmosphere[1]. Primary productivity over large areas of the EEP is limited by silicic acid and iron availability, and because of this constraint the organic carbon export to the deep ocean is unable to compensate for the outgassing of carbon dioxide that occurs through upwelling of deep waters. It has been suggested that the delivery of dust-borne iron to the glacial ocean[2,3] could have increased primary productivity and enhanced deep-sea carbon export in this region, lowering atmospheric carbon dioxide concentrations during glacial periods. Such a role for the EEP is supported by higher organic carbon burial rates documented in underlying glacial sediments[4,5], but lower opal accumulation rates cast doubts on the importance of the EEP as an oceanic region for significant glacial carbon dioxide drawdown[6,7]. Here we present a new silicon isotope record that suggests the paradoxical decline in opal accumulation rate in the glacial EEP results from a decrease in the silicon to carbon uptake ratio of diatoms under conditions of increased iron availability from enhanced dust input. Consequently, our study supports the idea of an invigorated biological pump in this region during the last glacial period that could have contributed to glacial carbon dioxide drawdown. Additionally, using evidence from silicon and nitrogen isotope changes, we infer that, in contrast to the modern situation, the biological productivity in this region is not constrained by the availability of iron, silicon and nitrogen during the glacial period. We hypothesize that an invigorated biological carbon dioxide pump constrained perhaps only by phosphorus limitation was a more common occurrence in low-latitude areas of the glacial ocean.

The EEP is an important area of biogenic opal production and burial in the ocean[8], accounting for two-thirds of the marine carbon dioxide efflux to the atmosphere today[1]. This is because primary productivity in the EEP is co-limited by silicic acid and iron availability[9] and so the biological carbon dioxide pump is unable to compensate for the upwelling of carbon dioxide[10] (Fig. 1).

In Fig. 2, we present glacial–interglacial records of biogenic accumulation as well as stable nitrogen and silicon isotope records spanning the past 35,000 years from an EEP marine sediment Ocean Drilling Program (ODP) Core 1240 (Fig. 1). Organic carbon concentrations in the sediment during the Last Glacial Maximum (LGM) were twice those of the Holocene and peaked during the deglacial period. The ^{230}Th-normalized organic carbon accumulation rates also show the same trend (Fig. 2). The doubling of organic carbon accumulation during the LGM is in agreement with previous studies showing that organic carbon export in the EEP was higher during the LGM[4,5] and may imply an invigorated biological carbon dioxide pump during this period (see Supplementary Information). Contradicting such an assertion, opal concentration in Core 1240 was reduced roughly by a factor of two during the LGM, consistent with existing opal accumulation records in the region[6,7]. Such differences in biogenic accumulation records cannot be explained by local changes in upwelling alone.

The sedimentary ratio of silicon to organic carbon (Si:C) in Core 1240 increases drastically during the glacial to interglacial transition with values during the early Holocene almost three times higher than during the LGM (Fig. 2). This change in the relative accumulation of opal and carbon is seen across the entire EEP and therefore cannot be explained simply by factors such as differential distribution and preservation of these biogenic components (Supplementary Information).

The contrast between opal and organic carbon accumulation trends during the LGM in the EEP could be reconciled by invoking a shift in the phytoplankton communities from diatomaceous to coccolithophorid production. This would have reduced the carbon rain-rate ratios (the relative contributions of organic carbon and carbonate to total export production) in the glacial EEP and hence the capacity of the marine carbon reservoir to sequester carbon[11]. However, carbonate contents and accumulation rates in Core 1240 do not show a clear glacial–interglacial pattern to support such a shift

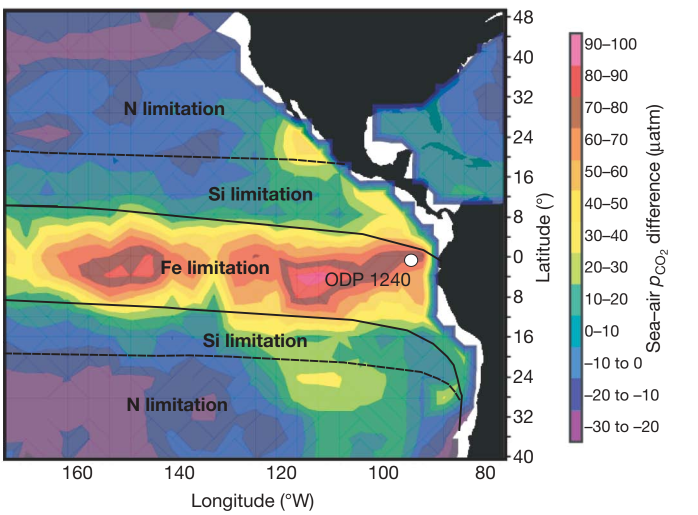


Figure 1 | Annually averaged carbon dioxide partial pressure (p_{CO_2}) difference between the atmosphere and the ocean for 1995. Data modified from ref. 1; see http://www.ldeo.columbia.edu/res/pi/CO2/. Also depicted is the hierarchy of nutrient limitation for diatom growth across the low-latitude Pacific Ocean. Black lines delimit the area where the most severe limitation is imposed by iron, dotted lines delimit the area of silicic acid limitation. Beyond these lines the ocean is first nitrate-limiting (modified from ref. 9). We note the geographical correspondence between carbon dioxide efflux maxima (warm colours) and iron and silicic acid limitation areas. The site of ODP 202 Core 1240 is shown with a white dot.

[1]School of Geosciences, Grant Institute, University of Edinburgh, West Main Road, EH10 3JW, Edinburgh, UK. [2]IGMR, ETH Zürich, Clausiusstrasse 25, CH-8092 Zürich, Switzerland. [3]GRC Geociències Marines, Facultat de Geologia, Universitat de Barcelona C/ Martí Franques s/n 08028 Barcelona, Spain. [4]Scottish Universities Environment Research Centre, Rankine Avenue, Scottish Enterprise Technology Park, East Kilbride, G75 0QF, UK.

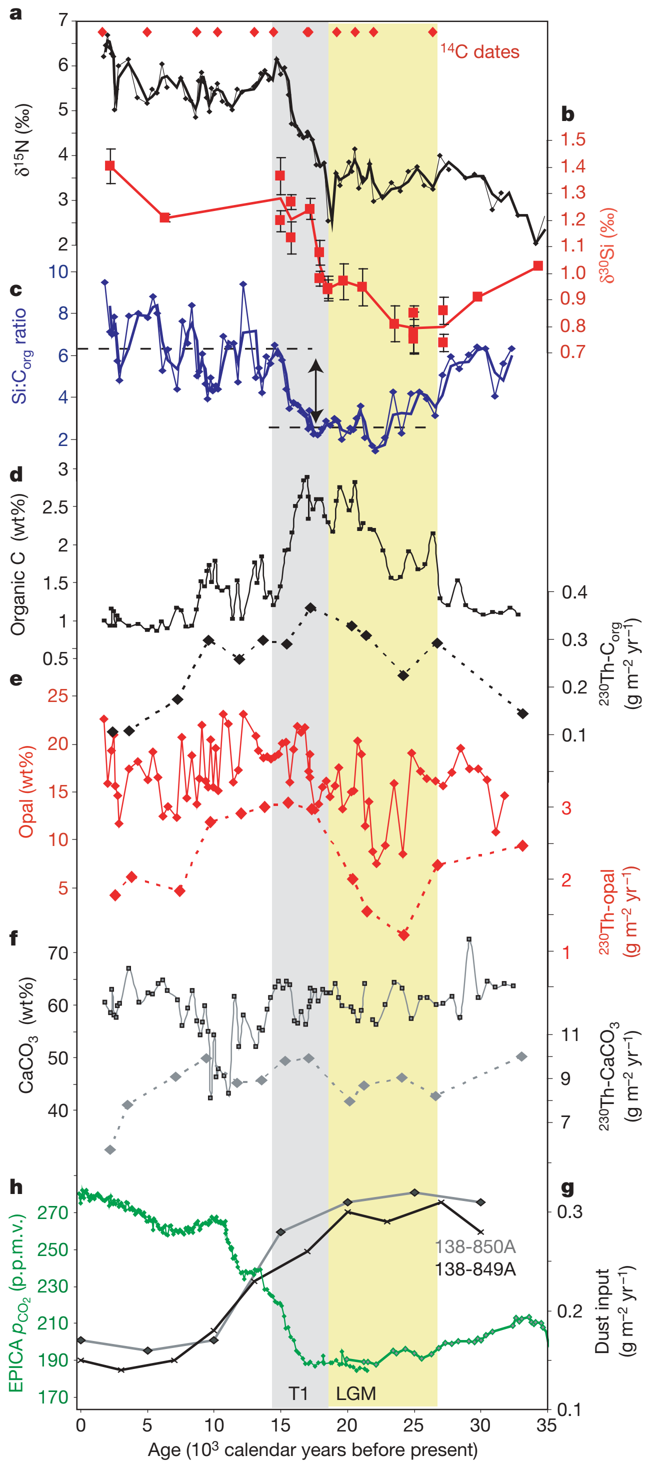


Figure 2 | Sedimentary records from Core 1240 plotted versus calendar ages before present. a, δ^{15}N signal of bulk material (the bold line indicates the 2-point average). **b,** The δ^{30}Si signal of diatoms (error bars are 1 sigma error of the mean). **c,** Si:C ratio (the bold line indicates the 2-point average). **d–f,** Elemental concentrations (solid lines) and ^{230}Th-normalized accumulation rates (dotted lines) of organic carbon (**d**), opal (**e**) and carbonate (**f**). **g,** Dust fluxes in Core ODP 138-849A and 138-850A (ref. 2). **h,** Atmospheric p_{CO_2} record from the EPICA (European Project for Ice Coring in Antarctica) dome C ice core[31] in p.p.m.v., parts per million by volume. T1, Termination 1.

in the rain-rate ratio of particulate export (Fig. 2). The estimates of carbonate accumulation rates in the EEP based either on the ^{230}Th-normalization technique or on high-resolution ^{14}C stratigraphy unequivocally document reduced carbonate fluxes and production during the LGM[12]. Moreover, a recent study[13] of carbonate accumulation corrected for sedimentary dissolution has even suggested that primary production in the EEP shifted from carbonate to opal production during the LGM compared to the Holocene. Therefore, the sedimentary opal records from the EEP showing halved accumulation rates during the LGM compared to the Holocene epoch[7], thus implying a less efficient biological carbon dioxide pump, contradict the inferences based on enhanced organic carbon production and have puzzled researchers. Resolving this paradox has important implications in understanding the role of the EEP in glacial carbon dioxide drawdown and the impact of enhanced iron delivery on nutrient limitation and the biological carbon dioxide pump during the glacial periods.

Availability of silicic acid in the surface water of the EEP is considered the major control on opal production[10]—a view that stems from modern observations of silicon limitation in this region (Fig. 1). However, LGM conditions could have been drastically different for diatom growth when dust-borne iron delivery to the surface EEP was twice as much as it is today, as recorded by sediment cores in this area[2,3] (Fig. 2). The impact of this increased iron delivery on opal production in the iron-limited low-latitude Pacific Ocean[2,3] has never been fully explored. Results from *in situ* iron-fertilization experiments[14,15] reveal that the alleviation of iron limitation in the EEP causes a two- to threefold decrease in Si:C uptake ratios of siliceous producers, owing to decreased cell silicification[14,15]. If this were to happen during the iron-replete LGM, lower silicon uptake during diatom growth may have resulted and consequently reduced opal accumulation. In addition, given the large decline in Si:C uptake ratios documented in these experiments, surplus silicic acid may have been available during diatom production. Thus, the reconstruction of glacial silicic acid utilization in the EEP could allow us to unravel the opal accumulation paradox and to resolve the carbon mass balance during the LGM.

We evaluate the possibility of lowered silicon utilization during diatom growth under iron-replete glacial periods using a record of δ^{30}Si in opal (Fig. 2). During the Holocene δ^{30}Si values are about +1.3‰, a value close to the isotopic composition of silicic acid in intermediate water masses of the tropical Pacific[16,17]. This indicates near-complete utilization of the available silicic acid, consistent with modern observations of its limitation in the EEP (Supplementary Information). In contrast, the LGM values were much lower (<0.9‰). This lower δ^{30}Si signature may reflect a decrease in the relative utilization of silicic acid in the EEP during the glacial period, provided this shift is not affected by glacial change in the nutrient source.

During the glacial period, silicic acid leakage from the Southern Ocean, whereby enhanced dust-borne iron input to the Southern Ocean resulted in the conservation of silicic acid over other nutrients, leading to its export to low latitudes[18,19], is one mechanism that could have potentially altered silicic acid supply to the EEP and its isotopic signature. It has been estimated that during the glacial period, the low-latitude Pacific could have received three times more silicic acid through the subantarctic mode waters, with δ^{30}Si values ranging from 1.7‰ to 3‰[17,20]. The resulting change in source signature can at best account only for a very modest decline in the glacial diatom-bound silicon isotope signal (~0.1‰, assuming the minimum value of 1.7‰ for the Southern Ocean source). However, given the range of values for the subantarctic source, heavier shifts in glacial sediments are more probable (see Supplementary Information). Additionally, the threefold increase in the supply of $Si(OH)_4$ to the EEP should have increased glacial opal burial. On the contrary, Core 1240 records a glacial decline in opal accumulation with markedly lighter isotopic values (Fig. 2).

The reason for the apparent contrast becomes clear if we consider how the glacial increase in dust-borne iron inputs would have affected silicic acid utilization locally within the EEP. The increased iron input and the resultant decline in Si:C uptake ratios, as proposed for the

Southern Ocean, should also have operated in the glacial EEP, leading to the conservation of silicic acid[14,18]. This would explain the decline in opal accumulation and the lighter $\delta^{30}Si$ values in the glacial intervals of Core 1240. Paradoxically, the increased silicic acid supply from the Southern Ocean appears to have occurred at the time the demand for this nutrient declined locally within the EEP—a contention that is supported by the similar glacial–interglacial histories of dust input in these two regions[3]. Although regional patterns in glacial silicic acid utilization need further examination, this new result documenting excess silicic acid in the glacial EEP strongly implies that the decline in opal accumulation in LGM sediments is not controlled by the availability of silicic acid but is most probably the consequence of iron fertilization and lowered Si:C uptake ratios during diatom growth. Importantly, the glacial decrease in opal accumulation observed across the EEP[6,7] does not reflect a decline in carbon rain-rate ratio and the biological carbon dioxide pump, as was previously suggested[6].

Our results suggest that the glacial scenario of nutrient limitation and the constraints on biological production in the EEP were drastically different from the modern situation depicted in Fig. 1. In contrast to modern conditions, enhanced iron delivery and the resultant switch to lower Si:C uptake ratio by diatoms led to conditions where iron and silicon are no longer limiting factors for biological productivity. From glacial $\delta^{30}Si$ values in Core 1240, and applying a steady-state model[17] we estimate that relative silicic acid utilization during the LGM declined by about 25% from a near-complete utilization during the Holocene. This calculation assumes a constant $Si(OH)_4$ source of around +1.2‰, in agreement with low-latitude Pacific intermediate-water values[16,17]. This estimate of decreased $Si(OH)_4$ utilization would be even greater if the leakage of isotopically heavy silicic acid from subantarctic waters is taken into account[20]. Similarly, the nitrogen-isotope record allows us to assess the potential for glacial nitrate limitation. The absence of iron and silicon co-limitation during the glacial period should have led to increased nitrate utilization and heavier glacial $\delta^{15}N$.

In contrast, the N-isotope profile (Fig. 2) shows a trend towards lower $\delta^{15}N$ values during the LGM relative to the Holocene, consistent with other, similar studies in this region[21]. However, given that nitrate utilization is incomplete in the modern EEP, a further $\delta^{15}N$ decrease (>3‰) during the glacial period cannot be explained simply by changes in nitrate uptake by biota alone because this would entail a very large reduction in nitrate utilization. Therefore, we suggest that the low LGM $\delta^{15}N$ values relative to the Holocene reflect an additional supply of isotopically lighter nitrogen due to reduced denitrification in the oxygen minimum zones off Peru and Mexico[22,23]. Today, the prevalence of denitrification in these margins bordering the EEP means that the source waters supplied to the EEP thermocline are depleted in nitrate relative to phosphate (lower than the Redfield N:P value) and have a relatively heavy nitrogen-isotope signature (+6.5‰ as opposed to the +4.8‰ ocean average) owing to mixing with partially denitrified heavy nitrate[24]. Thus, reduced denitrification during glacial periods would enhance the nitrate inputs into the EEP by increasing the N:P ratio of upwelling waters and provide nitrate that is isotopically lighter. In combination, both these factors would lead to a shift to lighter sedimentary nitrogen-isotope signatures irrespective of any glacial change in local upwelling[25,26]. Thus, the lighter N-isotope values provide evidence that nitrate is also not a limiting nutrient in the glacial EEP—a condition comparable to the modern situation. Therefore, the data strongly suggest simultaneous removal of the constraints imposed by iron, silicon and nitrogen limitation on biological production in the EEP during the LGM.

This argues for a more invigorated biological carbon dioxide pump in the glacial EEP, perhaps ultimately constrained by the availability of dissolved phosphate. The resulting increase in carbon export and rain-rate ratio as documented in the glacial sediments of the EEP (Supplementary Information) should have reduced carbon dioxide evasion. Such an assertion is also consistent with the timing of the first phase of the atmospheric carbon dioxide rise during the last glacial termination (~18,000 years ago) which is contemporaneous (within dating uncertainties) with the decline in dust delivery and the increases in sedimentary Si:C and $\delta^{30}Si$ in EEP records.

Finally, we suggest that our results have much wider ramifications: the EEP provides an illustration of the constraints imposed on the biological carbon dioxide pump in low-latitude oceans of the glacial periods. In currently iron-limited open-ocean 'high-nitrate, low-chlorophyll' regions, such as the EEP and the Southern Ocean, the switch to iron-replete LGM conditions has led to the generation of excess silicic acid during diatom growth, which in turn is subject to dispersal through surface and subsurface ocean circulation, increasing the availability of silicic acid over a much wider area of the glacial ocean[18,19]. This additional silicic acid supply occurs at a time when its demand by biota is already reduced by iron fertilization, which should have caused drastic reductions in silicon limitation over large areas of the glacial ocean.

Also, the documented glacial decline in denitrification in the world's oxygen minimum zones is expected to have increased the availability of nitrate[23,27]. For instance, the NO_3:PO_4 ratio at the surface of the modern EEP is about 12.5 (that is, lower than the Redfield N:P ratio of 16). In contrast, a 30% increase in nitrate inventory during the last glacial period compared to today, as suggested by modelling work[28], could have been sufficient to cause the N:P ratio in the EEP to exceed the Redfield ratio. Such changes, occurring more widely, would have resulted in an invigorated carbon dioxide pump in low-latitude oceans ultimately constrained by the availability of dissolved phosphate, as suggested by the case study reported here and predicted by recent model results[29]. Therefore, we hypothesize that phosphorus limitation was much more widespread during the glacial periods—a situation fundamentally different to that of the modern ocean.

METHODS SUMMARY

ODP 202 Core 1240 was retrieved from the Cocos ridge (00°01.31′ N; 86°27.76′ W; 2,921 m depth) in the EEP. The age model is based on 13 accelerator mass spectrometry (AMS) ^{14}C dates on planktonic foraminifers[25]. Determination of the opal content (wt%) was performed by molybdate-blue spectrophotometry on alkaline extracts. Organic carbon (wt%), total nitrogen (wt%) and $\delta^{15}N$ (wt‰) were determined on 10 mg of freeze-dried and powdered bulk sediments using a Carlo Erba elemental analyser coupled to a VG Prism III mass spectrometer at the University of Edinburgh. ^{230}Th normalization was performed using acid digestion on bulk samples and column chemistry followed by multiple-collector inductively coupled plasma mass spectrometry (MC-ICPMS) analyses. Purification of the diatom samples for silicon isotope measurement was performed by chemical leaching of the carbonate and organic fractions, sieving and differential settling. Silicon isotope determination was conducted in ETH Zürich on the Nu1700 high-resolution MC-ICPMS[30].

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 1 April 2008; accepted 23 April 2009.

1. Takahashi, T. *et al.* Global sea-air CO_2 flux based on climatological surface ocean pCO_2, and seasonal biological and temperature effects. *Deep-Sea Res. II* **49**, 1601–1622 (2002).
2. McGee, D., Marcantonio, F. & Lynch-Stieglitz, J. Deglacial changes in dust flux in the eastern equatorial Pacific. *Earth Planet. Sci. Lett.* **257**, 215–230 (2007).
3. Winckler, G., Anderson, R. F., Fleisher, M. Q., McGee, D. & Mahowald, N. Covariant glacial-interglacial dust fluxes in the equatorial Pacific and Antarctica. *Science* **320**, 93–96 (2008).
4. Pedersen, T. F. Increased productivity in the eastern equatorial Pacific during the Last Glacial Maximum (19,000 to 14,000 yr BP). *Geology* **11**, 16–19 (1983).
5. Lyle, M. Climatically forced organic-carbon burial in equatorial Atlantic and Pacific oceans. *Nature* **335**, 529–532 (1988).
6. Kienast, S. S., Kienast, M., Jaccard, S., Calvert, S. E. & Francois, R. Testing the silica leakage hypothesis with sedimentary opal records from the eastern equatorial Pacific over the last 150 kyrs. *Geophys. Res. Lett.* **33**, doi:10.1029/2006GL026651 (2006).
7. Bradtmiller, L. I., Anderson, R. F., Fleisher, M. Q. & Burckle, L. H. Diatom productivity in the equatorial Pacific Ocean from the last glacial period to the present: A test of the silicic acid leakage hypothesis. *Paleoceanography* **21**, doi:10.1029/2006PA001282 (2006).
8. Lisitzin, A. P. in *The Micropalaeontology of Oceans Symposium, Illustrated Maps* 173–195 (Cambridge University Press, 1971).

9. Moore, J. K., Doney, S. C. & Lindsay, K. Upper ocean ecosystem dynamics and iron cycling in a global three-dimensional model. *Glob. Biogeochem. Cycles* **18**, doi:10.1029/2004GB002220 (2004).
10. Dugdale, R. C. & Wilkerson, F. P. Silicate regulation of new production in the equatorial Pacific upwelling. *Nature* **391**, 270–273 (1998).
11. Archer, D. & Maier-Reimer, E. Effect of deep-sea sedimentary calcite preservation on atmospheric CO_2 concentration. *Nature* **367**, 260–263 (1994).
12. Loubere, P. & Richaud, M. Some reconciliation of glacial-interglacial calcite flux reconstructions for the eastern equatorial Pacific. *Geochem. Geophys. Geosyst.* **8**, doi:10.1029/2006GC001367 (2007).
13. Richaud, M., Loubere, P., Pichat, S. & Francois, R. Changes in opal flux and the rain ratio during the last 50,000 years in the equatorial Pacific. *Deep-Sea Res. II* **54**, 762–771 (2007).
14. Takeda, S. Influence of iron availability on nutrient consumption ratio of diatoms in oceanic waters. *Nature* **393**, 774–777 (1998).
15. Hutchins, D. A. & Bruland, K. W. Iron-limited diatom growth and Si: N uptake ratios in a coastal upwelling regime. *Nature* **393**, 561–564 (1998).
16. Reynolds, B. C., Frank, M. & Halliday, A. N. Silicon isotope fractionation during nutrient utilization in the North Pacific. *Earth Planet. Sci. Lett.* **244**, 431–443 (2006).
17. Beucher, C. P., Brzezinski, M. A. & Jones, J. L. Sources and biological fractionation of silicon isotopes in the eastern equatorial Pacific. *Geochim. Cosmochim. Acta* **72**, 3063–3073 (2008).
18. Brzezinski, M. A. *et al.* A switch from $Si(OH)_4$ to NO_3 depletion in the glacial Southern Ocean. *Geophys. Res. Lett.* **29**, doi:10.1029/2001GL014349 (2002).
19. Matsumoto, K., Sarmiento, J. L. & Brzezinski, M. A. Silicic acid leakage from the Southern Ocean: A possible explanation for glacial atmospheric pCO_2. *Glob. Biogeochem. Cycles* **16**, doi:10.1029/2001GB001442 (2002).
20. Beucher, C. P., Brzezinski, M. A. & Crosta, X. Silicic acid dynamics in the glacial sub-Antarctic: implications for the silicic acid leakage hypothesis. *Glob. Biogeochem. Cycles* **21**, doi:10.1029/2006GB002746 (2007).
21. Farrell, J. W., Pedersen, T. F., Calvert, S. E. & Nielsen, B. Glacial–interglacial changes in nutrient utilization in the equatorial Pacific Ocean. *Nature* **377**, 514–517 (1995).
22. Toggweiler, J. R. & Carson, S. in *Upwelling in the Ocean: Modern Processes and Ancient Records* (eds Summerhayes, C. P., Emeis, K.-C., Angel, M. V., Smith, R. L. & Zeitzschel, B.) 337–360 (John Wiley & Sons, 1995).
23. Ganeshram, R. S., Pedersen, T. F., Calvert, S. E., McNeill, G. W. & Fontugne, M. R. Glacial-interglacial variability in denitrification in the world's oceans: causes and consequences. *Paleoceanography* **15**, 361–376 (2000).
24. Altabet, M. A. Nitrogen isotopic evidence for micronutrient control of fractional NO3- utilization in the equatorial Pacific. *Limnol. Oceanogr.* **46**, 368–380 (2001).
25. Pena, L. D., Cacho, I., Ferretti, P. & Hall, M. A. El Niño–Southern Oscillation–like variability during glacial terminations and interlatitudinal teleconnections. *Paleoceanography* **23**, doi:10.1029/2008PA001620 (2008).
26. Koutavas, A., Lynch-Stieglitz, J., Marchitto, T. M. & Sachs, J. P. El Nino-like pattern in ice age tropical Pacific sea surface temperature. *Science* **297**, 226–230 (2002).
27. Ganeshram, R. S., Pedersen, T. F., Calvert, S. E. & Francois, R. Reduced nitrogen fixation in the glacial ocean inferred from changes in marine nitrogen and phosphorus inventories. *Nature* **415**, 156–159 (2002).
28. Deutsch, C., Sigman, D. M., Thunell, R. C., Meckler, A. N. & Haug, G. H. Isotopic constraints on glacial/interglacial changes in the oceanic nitrogen budget. *Glob. Biogeochem. Cycles* **18**, doi:10.1029/2003GB002189 (2004).
29. Moore, J. K. & Doney, S. C. Iron availability limits the ocean nitrogen inventory stabilizing feedbacks between marine denitrification and nitrogen fixation. *Glob. Biogeochem. Cycles* **21**, doi:10.1029/2006GB002762 (2007).
30. Georg, R. B., Reynolds, B. C., Frank, M. & Halliday, A. N. New sample preparation techniques for the determination of Si isotopic compositions using MC-ICPMS. *Chem. Geol.* **235**, 95–104 (2006).
31. Monnin, E. *et al.* Atmospheric CO_2 concentrations over the last glacial termination. *Science* **291**, 112–114 (2001).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank R. Francois, S. Kienast, W. Geibert, E. Galbraith and G. Leduc for discussions. Funding for this project was provided by the Scottish Alliance for Geoscience Environment Society (SAGES), the Natural Environment Research Council (NERC) through a Standard NERC grant (to R.S.G. and L.E.P.) and a Marie Curie Intra-European fellowship (to L.E.P.). I.C. and L.P. acknowledge funding from the ROMIAT and GRACCIE-CONSOLIDER projects.

Author Contributions L.E.P. and R.S.G. initiated the project. L.E.P. and B.C.R. measured silicon isotopes in ETH Zurich, L.E.P. measured elemental composition and nitrogen isotopes, L.E.P., K.K. and R.M.E. made the ^{230}Th measurements. I.C. and L.P. provided the age model. L.E.P. and R.S.G. wrote the paper with the participation of B.R.C. All authors participated in the discussions on the results and commented on the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to L.E.P. (laetitia.pichevin@ed.ac.uk).

METHODS

ODP 202 Core 1240 was retrieved from the Cocos ridge (see ODP report http://www-odp.tamu.edu/publications/202_IR/chap_11/chap_11.htm).

Determination of the opal content (wt%) was performed according to the method published in ref. 32. ^{230}Th normalization was used to account for the effect of sediment focusing[33] and performed at the Scottish Universities Environment Research Centre[34] (Supplementary Information).

Purification of the diatom samples for silicon isotope measurement was performed following a cleaning method published recently[35]. A small amount of biogenic opal (0.5 mg) was dried down with concentrated perchloric acid at $\sim$180 °C in Teflon vials, and then dissolved in 100 µl of 1 M NaOH, before being diluted to 5 ml with 0.01 M HCl after 24 h. An equivalent to 10 µg of opal was loaded onto a pre-cleaned 1.8 ml DOWEX 50W-X12 cation exchange resin bed (in H^+ form) and eluted with 5 ml of purified water (Milli-Q element 18.2 MΩ cm^{-1}). The silicon isotope composition was determined on the diluted solution (0.6 p.p.m. Si) on the Nu1700 high-resolution MC-ICPMS at ETH Zürich, using a standard-sample-standard bracketing protocol. All results in this study were calculated using the δ^{30}Si notation for deviations of the measured ^{30}Si/^{28}Si from the international silicon standard NBS28 in parts per thousand (‰). The long-term reproducibility was better than 0.07‰ δ^{30}Si (one standard deviation, s.d.)[36]. Samples were measured at least five times, which resulted in a 95% confidence level below 0.08‰. Error bars on the δ^{30}Si plot are calculated as 1 sigma error of the mean (Fig. 2).

32. Mortlock, R. A. & Froelich, P. N. A simple method for the rapid determination of biogenic opal in pelagic marine sediments. *Deep-Sea Res. A* **36,** 1415–1426 (1989).
33. Francois, R., Frank, M., van der Loeff, M. M. R. & Bacon, M. P. Th-230 normalization: an essential tool for interpreting sedimentary fluxes during the late Quaternary. *Paleoceanography* **19,** doi:10.1029/2003PA000939 (2004).
34. Ellam, R. M. & Keefe, K. MC-ICP-MS analysis of non-natural U isotope ratios using a Th-229/Th-232 external mass bias correction. *J. Anal. At. Spectrom.* **22,** 147–152 (2007).
35. Morley, D. W. *et al.* Cleaning of lake sediment samples for diatom oxygen isotope analysis. *J. Paleolimnol.* **31,** 391–401 (2004).
36. Reynolds, B. C., Georg, R. B., Oberli, F., Wiechert, U. H. & Halliday, A. N. Re-assessment of silicon isotope reference materials using high-resolution multi-collector ICP-MS. *J. Anal. At. Spectrom.* **21,** 266–269 (2006).

Vol 459 | 25 June 2009 | doi:10.1038/nature08089

LETTERS

^{142}Nd evidence for an enriched Hadean reservoir in cratonic roots

Dewashish Upadhyay[1], Erik E. Scherer[1] & Klaus Mezger[1]

The isotope ^{146}Sm undergoes α-decay to ^{142}Nd, with a half-life of 103 million years. Measurable variations in the ^{142}Nd/^{144}Nd values of rocks resulting from Sm–Nd fractionation could therefore only have been produced within about 400 million years of the Solar System's formation (that is, when ^{146}Sm was extant). The ^{142}Nd/^{144}Nd compositions of terrestrial rocks[1–7] are accordingly a sensitive monitor of the main silicate differentiation events that took place in the early Earth. High ^{142}Nd/^{144}Nd values measured in some Archaean rocks from Greenland[1–5] hint at the existence of an early incompatible-element-depleted mantle. Here we present measurements of low ^{142}Nd/^{144}Nd values in 1.48-gigayear-(Gyr)-old lithospheric mantle-derived alkaline rocks from the Khariar nepheline syenite complex in southeastern India[8]. These data suggest that a reservoir that was relatively enriched in incompatible elements formed at least 4.2 Gyr ago and traces of its isotopic signature persisted within the lithospheric root of the Bastar craton until at least 1.48 Gyr ago. These low ^{142}Nd/^{144}Nd compositions may represent a diluted signature of a Hadean (4 to 4.57 Gyr ago) enriched reservoir that is characterized by even lower values. That no evidence of the early depleted mantle has been observed in rocks younger than 3.6 Gyr (refs 3, 4, 7) implies that such domains had effectively mixed back into the convecting mantle by then. In contrast, some early enriched components apparently escaped this fate. Thus, the mantle sampled by magmatism since 3.6 Gyr ago may be biased towards a depleted composition that would be balanced by relatively more enriched reservoirs that are 'hidden' in Hadean crust[6], the D′′ layer[9–11] of the lowermost mantle or, as we propose here, also within the roots of old cratons.

Major geochemical differentiation events such as the formation and crystallization of a magma ocean or the extraction of the first crust early in planetary history can fractionate Sm from Nd and produce reservoirs having contrasting ^{142}Nd/^{144}Nd compositions (see refs 1–7, 12–17). The variability of ^{142}Nd/^{144}Nd within terrestrial samples[1–7] and the range in ^{176}Hf/^{177}Hf of Hadean zircons from the Jack Hills (Australia)[18–20] provide unequivocal evidence that the early Earth underwent large-scale silicate differentiation during the Hadean, producing an early depleted mantle and an enriched reservoir (Fig. 1).

Here we report high-precision Nd isotope measurements on Mesoproterozoic alkaline rocks from southeastern India. Some of these have pronounced ^{142}Nd/^{144}Nd deficits relative to the La Jolla Nd standard (Supplementary Table S1). The source region of these rocks therefore contained the Nd isotopic signature of an incompatible-element-enriched component that formed in the Hadean. This is further evidence that the Earth underwent large-scale silicate differentiation early in its history. The discovery of rocks as young as 1.48 Gyr that have low ^{142}Nd/^{144}Nd values, and have thus sampled a relatively enriched Hadean source, is remarkable and indicates that evidence for the early differentiation of the Earth may be preserved in other geologic settings besides early Archaean terranes.

The samples come from the Khariar alkaline complex, which comprises a ~1.48-Gyr-old[8] differentiated magmatic suite dominated by nepheline syenite, syenite and minor ijolite, which are in places interlayered with tholeiites. The alkaline complex intruded into granite gneisses of the Archaean Bastar (or Bhandara) craton and is bordered by younger granulite facies rocks of the Eastern Ghats Belt to the east (Fig. 2). The cratonic basement gneisses consist of ~2.5-Gyr-old granitoids that include older (~3.5-Gyr-old) gneissic enclaves of a deformed tonalite–trondhjemite–granodiorite association. The late Mesoproterozoic (~1.1-Gyr-old) granulites bordering the alkaline complex formed during the collision of peninsular India with East Antarctica during the assembly of the supercontinent Rodinia[8]. Their earlier granulite facies mineral assemblages were overprinted by an amphibolite facies fabric during pan-African orogenesis (~0.5 Gyr ago), which deformed and metamorphosed the alkaline rocks as well[8].

The nepheline syenites and syenites are coarse-grained rocks consisting mostly of perthite, nepheline, and aligned grains of amphibole and biotite, with some of the latter having replaced clinopyroxene. Accessory phases include calcite, allanite, titanite, apatite, zircon, oxides (ilmenite and magnetite), and sulphides (pyrite, chalcopyrite, and pyrrhotite). Members of the alkaline suite are characterized by strong enrichment in the light-rare-earth elements and large-ion lithophile elements[8]. The most primitive rock type is a mesocratic variety of nepheline syenite. It is enriched in the light-rare-earth elements, U, Th and Nb, suggestive of rift-related magmatism. The geochemical variation in the suite has been explained by the differentiation of a primary basanitic magma derived from the lithospheric mantle root of the Bastar craton[8]. Relatively high initial ^{207}Pb/^{204}Pb (15.90–16.26) values[8] suggest that the alkaline magmas were derived from an old and incompatible-element-enriched source. The Khariar alkaline body forms a part of a linear chain of deformed alkaline complexes that delineate a collisional suture superimposed upon an earlier Mesoproterozoic (~1.5–1.3-Gyr-old) rift at the southeastern margin of the Indian shield (Fig. 2).

Four samples from the Khariar alkaline complex have well-resolved ^{142}Nd/^{144}Nd deficits of 8–13 parts per million (p.p.m.) relative to the La Jolla Nd standard. Three others have ^{142}Nd/^{144}Nd values that are identical to that of the terrestrial standard (Fig. 1). The anomalies are reproducible for repeated measurements of Nd from a single digestion as well as for multiple digestions of a single sample. The initial ε^{143}Nd values (referenced to the chondritic uniform reservoir) of all samples are negative (−0.25 to −4.39) at 1.48 Gyr ago. The fact that some samples have anomalous ^{142}Nd/^{144}Nd compositions whereas others do not at similar ^{143}Nd/^{144}Nd values indicates decoupling of the two isotope systems and variable preservation of the primary Nd isotope signature. Any mixing with Nd of 'normal' isotopic composition (that is, identical to the La Jolla Nd standard) that took place during alkaline magma genesis and differentiation or metamorphism

[1]Zentrallabor für Geochronologie, Institut für Mineralogie, Universität Münster, Corrensstrasse 24, 48149 Münster, Germany.

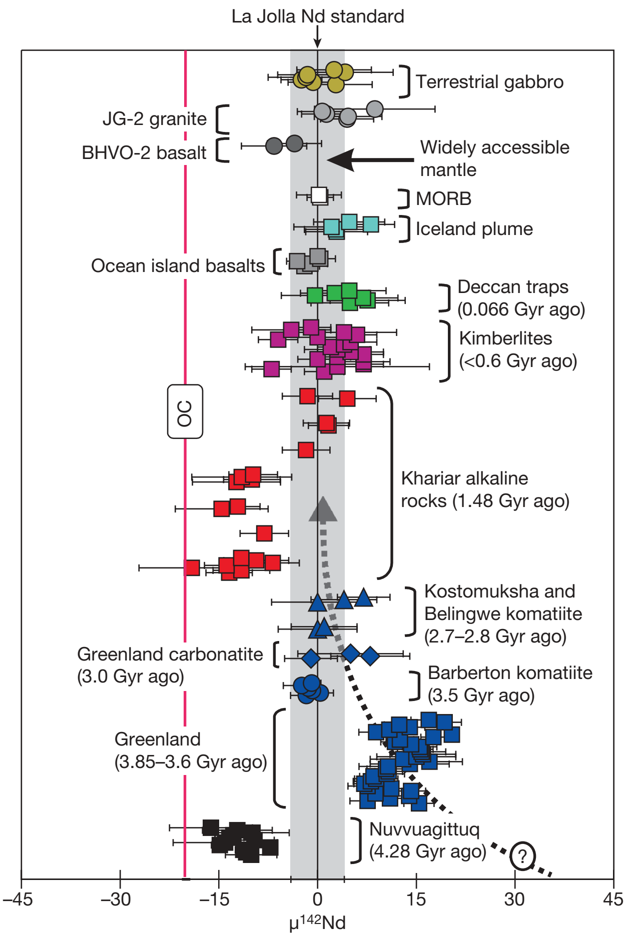


Figure 1 | $^{142}Nd/^{144}Nd$ of the Khariar alkaline rocks and other mantle-derived rocks[3–7]. Data are plotted as deviations in p.p.m. ($\mu^{142}Nd$) from the terrestrial Nd standard (vertical line with shaded grey bar as 2 s.d. uncertainty). Error bars are 2 standard errors (s.e.) of individual measurements. The alkaline rocks have well-resolved negative anomalies, indicating that they sampled a relatively enriched Hadean source. The data groups are arranged chronologically (not to scale) with the oldest rocks at the bottom (except for the BHVO2 and JG2 standards and a terrestrial gabbro). We note that there are no signs of the early depleted reservoir (that is, ^{142}Nd excesses) in any <3.6-Gyr-old mantle-derived rocks, suggesting that the early depleted material (dotted curve) had completely remixed into the accessible mantle by that time. In contrast, a component from an early, relatively enriched reservoir persisted until at least 1.48 Gyr ago, when it was sampled by the parent magmas of the Khariar alkaline rocks. If such preservation of enriched reservoirs is significant, the accessible mantle as sampled by magmatism since 3.6 Gyr ago must be biased towards a depleted composition. OC, average of ordinary chondrites; MORB, mid-ocean-ridge basalt.

would have shifted the $^{142}Nd/^{144}Nd$ values upwards towards that of the terrestrial standard. This implies that even larger negative anomalies existed and the measured values are just diluted remnants of these. The $^{142}Nd/^{144}Nd$ values of these rocks are similar to those in some rocks from the Nuvvuagittuq greenstone belt[6], the only other location on the Earth where such isotopic compositions have been measured. This signature complements the high $^{142}Nd/^{144}Nd$ values that now appear to be a common feature in the Archaean rocks of southwest Greenland[1–5]. Because the alkaline magmas of the Khariar complex did not assimilate significant amounts of crust during emplacement[8], we interpret the low $^{142}Nd/^{144}Nd$ in the nepheline syenites to be a primary feature inherited from a relatively enriched source that formed within a few hundred million years of the Earth's accretion. The Khariar alkaline magmas were derived from the lithospheric mantle[8], so traces of this ancient enriched signature must still have existed in the lithospheric root of the Bastar craton at 1.48 Gyr ago.

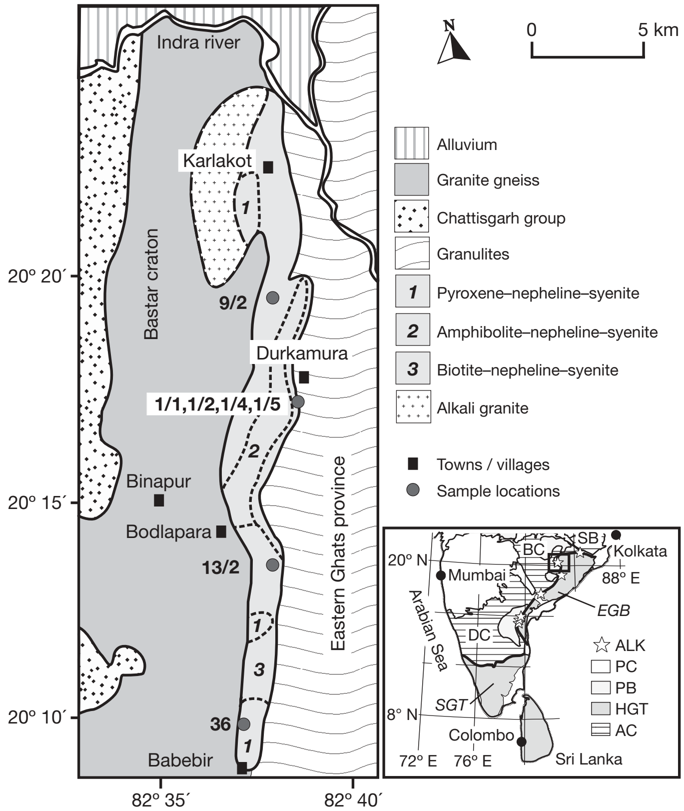


Figure 2 | Geological map of the Khariar alkaline complex showing sample locations. (The DU prefix in sample numbers has been omitted for clarity.) The inset shows peninsular India with the Archaean cratons and alkaline complexes at their southeastern margins. The Khariar alkaline body is marked by a rectangle. ALK, alkaline complexes; PC, Phanerozoic cover; PB, Proterozoic basins; HGT, high-grade terrane; AC, Archaean craton; SB, Singhbhum craton; BC, Bastar or Bhandara craton; DC, Dharwar craton; EGB, Eastern Ghats belt; SGT, Southern Granulite terrane. GPS coordinates of samples are as follows. DU-1/1, DU-1/2, DU-1/4 and DU-1/5: 20° 17.24′ N, 82° 38.48′ E; DU-13/2: 20° 13.37′ N, 82° 38.06′ E; DU-36: 20° 09.69′ N, 82°36.88′E; DU-9/2: 20° 19.67′ N, 82° 37.96′ E.

That an early enriched reservoir signature persisted until the Mesoproterozoic although no trace of the early depleted mantle has been identified in any rock younger than 3.6 Gyr (refs 3, 4, 7; Fig. 1) may indicate that the early depleted reservoir was part of the actively convecting mantle and hence was rapidly remixed with primitive or enriched components. In contrast, some early, relatively enriched reservoirs may have survived as pockets of Hadean crust, perhaps in the Nuvvuagittuq greenstone belt[6], or as mantle components either in the lithospheric roots of old continents (as we propose for the Khariar alkaline rock source) or in the D′′ layer of the lower mantle[9–11]. Seismic tomography and magnetotelluric measurements of the mantle under Archaean cratons show that the stable, cratonic parts of continents have cold, deep roots[21] that can remain fixed to the crust and thereby isolated from convective recycling since Archaean time[22]. Heterogeneities in the underlying mantle may become incorporated into continental roots as the root deepens via cooling or as the continental keel sweeps laterally through the mantle[23]. Because these domains do not participate in convection, they are suitable places for preserving early enriched reservoirs for long periods of time without completely erasing their Nd isotope signatures. Thus, in addition to the D′′ layer of the deep mantle[9–11], stable continents are the other locations where remnants of a distinct Hadean enriched reservoir could remain inaccessible to convection over the vast time spans (more than 2.7 Gyr) required.

Because the bulk silicate Earth (BSE) $^{142}Nd/^{144}Nd$ composition is currently unconstrained, $^{142}Nd/^{144}Nd$ anomalies of the Khariar

nepheline syenites cannot be unambiguously characterized as a depleted or an enriched mantle signature. The $^{142}Nd/^{144}Nd$ value of these rocks is higher than the mean of ordinary chondrites[9,24,25] by ~6 p.p.m. (Fig. 1). If the BSE consists of material similar to ordinary chondrites, as is generally assumed[2–6,9], the higher-than-chondritic $^{142}Nd/^{144}Nd$ composition of even the most unradiogenic Khariar and Nuvvuagittuq[6] samples would indicate derivation from an early depleted mantle. If the entire present-day mantle is assumed to be depleted[9], mass balance constraints (BSE = depleted mantle + crust + early enriched reservoir) would then require the presence of an untapped enriched reservoir having a $^{142}Nd/^{144}Nd$ value that is 70 to 260 p.p.m. lower than ordinary chondrites for depleted-mantle Nd concentrations between 0.80 and 1.0 p.p.m. (Supplementary Table S2). If the differentiation that produced these reservoirs took place 4.5 Gyr ago (refs 3, 9), the early enriched reservoir would need to evolve with a $^{147}Sm/^{144}Nd$ of between 0.006 and 0.145 to generate the aforementioned $^{142}Nd/^{144}Nd$ deviation from ordinary chondrites. The lower third of this range is petrologically unreasonable, that is, rocks having such compositions have not been commonly reported on the Earth because the similar partitioning behaviour of Sm and Nd prevents extreme fractionation between these elements during magmatic processes.

Recent studies of chondritic meteorites have reported considerable variation in $^{142}Nd/^{144}Nd$ composition (−7 to −45 p.p.m. relative to the terrestrial standard) that does not correlate with Sm/Nd[24,25]. This has been partially attributed to p- or s-process nucleosynthetic heterogeneity[24,25] among Solar System objects and opens up the possibility that the Earth and other terrestrial planets accreted from material that had a $^{142}Nd/^{144}Nd$ composition different from that of ordinary chondrites. This appears to be the case for Martian shergottites because their $^{142}Nd/^{144}Nd$ versus Sm/Nd correlation, which has been variably interpreted as an isochron and a mixing line, passes through a $^{142}Nd/^{144}Nd$ value of about +10 p.p.m. relative to ordinary chondrites at chondritic $^{147}Sm/^{144}Nd$[16,26]. If the BSE had a $^{142}Nd/^{144}Nd$ value ~10 p.p.m. higher than ordinary chondrites, mass balance constraints would then require the Hadean enriched reservoir to have a $^{142}Nd/^{144}Nd$ value that is 25 to 120 p.p.m. lower than ordinary chondrites for depleted-mantle Nd concentrations between 0.80 and 1.0 p.p.m. (see Supplementary Information). For differentiation at 4.5 Gyr ago, this reservoir would need a $^{147}Sm/^{144}Nd$ between 0.110 and 0.178 to generate these estimated deviations from ordinary chondrites. This entire range, in contrast to that of the previous example, is petrologically reasonable.

For either scenario (that is, either BSE = ordinary chondrites or BSE = ordinary chondrites + 10 p.p.m.), the $^{142}Nd/^{144}Nd$ composition of the source for the Khariar alkaline suite must represent only a diluted remnant of the more negative signature of the Hadean enriched reservoir. It is conceivable that the enriched Hadean component formed during the crystallization of the terrestrial magma ocean and thus may have been similar to the potassium-, rare-earth-element- and phosphorus-enriched rocks from the Moon[9]. Such an enriched reservoir would have had a mass equivalent to 9% of the present-day crust (Supplementary Table S3).

The regional extent and the absolute volume of this enriched Hadean component are at present unknown, but its discovery in Proterozoic rocks that intruded into the margin of an Archaean continental fragment provides strong evidence that some of it may have been preserved in old cratons. The Khariar $^{142}Nd/^{144}Nd$ data suggest that some of these enriched components, in contrast to the early depleted reservoir, were not remixed back into the convecting mantle by 3.6 Gyr ago, but instead persisted to at least 1.48 Gyr ago. If significant pockets of this ancient enriched material still exist, within or beneath old cratons or in the D′′ layer, then the $^{142}Nd/^{144}Nd$ of the accessible mantle (represented by the La Jolla terrestrial Nd standard) sampled by igneous rocks less than 3.6 Gyr old must be biased towards a depleted composition.

METHODS SUMMARY

Whole-rock samples were powdered, dissolved and processed through ion exchange chromatography to isolate Sm and Nd. Rare-earth elements were eluted as a group in 6 M HCl on 4 ml of Bio-Rad 50W-X8 resin followed by the separation of Sm and Nd using reverse-phase chromatography on columns packed with Teflon powder coated with bis-(2-ethylhexyl) phosphoric acid. The Nd was further purified by eluting with 0.21 M α-hydroxy isobutyric acid (α-HIBA, pH = 4.17) on 50W-X4 resin in ammonia form. A final clean-up step to remove the α-HIBA from the Nd was performed using 6 M HCl on columns containing 1.5 ml of Bio-Rad 50W-X8 resin. Total blanks were <100 pg and <61 pg for Nd and Sm respectively. The $^{142}Nd/^{144}Nd$ and $^{143}Nd/^{144}Nd$ isotope compositions were measured as Nd^{+} on a Triton thermal ionization mass spectrometer using a multidynamic measurement routine[27].

Analyses of the La Jolla Nd standard ($n = 16$) interspersed with the samples yielded values of 1.1418326 (±4.0 p.p.m., 2 standard deviations, s.d.) for the $^{142}Nd/^{144}Nd$, and 0.5118454 (±7.3 p.p.m., 2 s.d.) for the $^{143}Nd/^{144}Nd$. Seven measurements of four digestions of a terrestrial gabbro yielded values of 1.1418331 (±5.3 p.p.m., 2 s.d.) for $^{142}Nd/^{144}Nd$ and 0.5127639 (±5.7 p.p.m., 2 s.d.) for $^{143}Nd/^{144}Nd$, demonstrating that our sample measurements are reproducible at the precision of our standard measurements. This level of precision was confirmed by five measurements from two digestions of the JG2 granite standard, which yielded values of 1.1418371 (±6.4 p.p.m., 2 s.d.) for $^{142}Nd/^{144}Nd$, and 0.5122185 (±5.2 p.p.m., 2 s.d.) for $^{143}Nd/^{144}Nd$.

Received 4 December 2008; accepted 23 April 2009.

1. Harper, C. L. & Jacobsen, S. B. Evidence from coupled $^{147}Sm–^{143}Nd$ and $^{146}Sm–^{142}Nd$ systematics for very early (4.5 Gyr) differentiation of the Earth's mantle. *Nature* **360,** 728–732 (1992).
2. Caro, G., Bourdon, B., Birck, J.-L. & Moorbath, S. $^{146}Sm–^{142}Nd$ evidence for early differentiation of the Earth's mantle. *Nature* **423,** 428–432 (2003).
3. Caro, G., Bourdon, B., Birck, J.-L. & Moorbath, S. High-precision $^{142}Nd/^{144}Nd$ measurements in terrestrial rocks: constraints on the early differentiation of the Earth's mantle. *Geochim. Cosmochim. Acta* **70,** 164–191 (2006).
4. Boyet, M. & Carlson, R. W. A new geochemical model for the Earth's mantle inferred from $^{146}Sm–^{142}Nd$ systematics. *Earth Planet. Sci. Lett.* **250,** 254–268 (2006).
5. Bennett, V. C., Brandon, A. D. & Nutman, A. P. Coupled ^{142}Nd-^{143}Nd isotopic evidence for Hadean mantle dynamics. *Science* **318,** 1907–1910 (2007).
6. O'Neil, J., Carlson, R. W., Francis, D. & Stevenson, R. K. Neodymium-142 evidence for Hadean mafic crust. *Science* **321,** 1828–1831 (2008).
7. Andreasen, R., Sharma, M., Subbarao, K. V. & Viladkar, S. G. Where on Earth is the enriched Hadean reservoir? *Earth Planet. Sci. Lett.* **266,** 14–28 (2007).
8. Upadhyay, D., Raith, M. M., Mezger, K., Bhattacharya, A. & Kinny, P. D. Mesoproterozoic rifting and Pan-African continental collision in SE India: evidence from the Khariar alkaline complex. *Contrib. Mineral. Petrol.* **151,** 434–456 (2006).
9. Boyet, M. & Carlson, R. W. ^{142}Nd evidence for early (4.53 Ga) global differentiation of the silicate Earth. *Science* **309,** 576–581 (2005).
10. Tolstikhin, I. & Hofmann, A. W. Early crust on top of the Earth's core. *Phys. Earth Planet. Inter.* **148,** 109–130 (2005).
11. Labrosse, S., Hernlund, J. W. & Coltice, N. A crystallizing dense magma ocean at the base of the Earth's mantle. *Nature* **450,** 866–869 (2007).
12. Harper, C. L. Jr, Nyquist, L. E., Bansal, B., Wiesmann, H. & Shih, C.-Y. Rapid accretion and early differentiation of Mars indicated by $^{142}Nd/^{144}Nd$ in SNC meteorites. *Science* **267,** 213–217 (1995).
13. Nyquist, L. E. *et al.* $^{146}Sm–^{142}Nd$ formation interval for the lunar mantle. *Geochim. Cosmochim. Acta* **13,** 2817–2837 (1995).
14. Foley, C. N. *et al.* The early differentiation history of Mars from ^{182}W-^{142}Nd isotope systematics in the SNC meteorites. *Geochim. Cosmochim. Acta* **69,** 4557–4571 (2005).
15. Rankenburg, K., Brandon, A. D. & Neal, C. R. Neodymium isotope evidence for a chondritic composition of the Moon. *Science* **312,** 1369–1372 (2006).
16. Debaille, V., Brandon, A. D., Yin, Q. Z. & Jacobsen, B. Coupled $^{142}Nd–^{143}Nd$ evidence for a protracted magma ocean in Mars. *Nature* **450,** 525–528 (2007).
17. Boyet, M. & Carlson, R. W. A highly depleted moon or a non-magma ocean origin for the lunar crust? *Earth Planet. Sci. Lett.* **262,** 505–516 (2007).
18. Amelin, Y. V., Lee, D. C., Halliday, A. N. & Pidgeon, R. T. Nature of the Earth's earliest crust from hafnium isotopes in single detrital zircons. *Nature* **399,** 252–255 (1999).
19. Harrison, T. M. *et al.* Heterogeneous Hadean hafnium: evidence of continental crust by 4.4–4.5 Ga. *Science* **310,** 1947–1950 (2005).
20. Harrison, T. M., Schmitt, A. K., McCulloch, M. T. & Lovera, O. M. Early (≥4.5 Ga) formation of terrestrial crust: Lu–Hf, $\delta^{18}O$, and Ti thermometry results for Hadean zircons. *Earth Planet. Sci. Lett.* **268,** 476–486 (2008).
21. Jordan, T. H. Continental tectosphere. *Rev. Geophys.* **13,** 1–12 (1975).
22. Maréschal, M. *et al.* Archaean cratonic roots, mantle shear zones and deep electrical anisotropy. *Nature* **375,** 134–137 (1995).
23. Helffrich, G. Heterogeneity in the mantle—its creation, evolution and destruction. *Tectonophysics* **416,** 23–31 (2006).
24. Andreasen, R. & Sharma, M. Solar nebula heterogeneity in p-process samarium and neodymium isotopes. *Science* **314,** 806–809 (2006).

25. Carlson, R. W., Boyet, M. & Horan, M. Chondrite barium, neodymium and samarium isotopic heterogeneity and early Earth differentiation. *Science* **316,** 1175–1178 (2007).
26. Caro, G., Bourdon, B., Halliday, A. N. & Quitté, G. Super-chondritic Sm/Nd ratios in Mars, the Earth and the Moon. *Nature* **452,** 336–339 (2008).
27. Upadhyay, D., Scherer, E. E. & Mezger, K. Fractionation and mixing of Nd isotopes during thermal ionization mass spectrometry: implications for high precision $^{142}Nd/^{144}Nd$ analyses. *J. Anal. At. Spectrom.* **23,** 561–568 (2008).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements D.U. thanks A. Bhattacharya and M. Raith for help during field work and sample collection. This project was funded by grants from the Deutsche Forschungsgemeinschaft (ME1717).

Author Contributions D.U. collected the samples, processed them through ion exchange chromatography and measured the isotopic compositions. All authors contributed to the writing and editing of the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to D.U. (upadhyay@uni-muenster.de).

LETTERS

Origins and evolutionary genomics of the 2009 swine-origin H1N1 influenza A epidemic

Gavin J. D. Smith[1], Dhanasekaran Vijaykrishna[1], Justin Bahl[1], Samantha J. Lycett[2], Michael Worobey[3], Oliver G. Pybus[4], Siu Kit Ma[1], Chung Lam Cheung[1], Jayna Raghwani[2], Samir Bhatt[4], J. S. Malik Peiris[1], Yi Guan[1] & Andrew Rambaut[2]

In March and early April 2009, a new swine-origin influenza A (H1N1) virus (S-OIV) emerged in Mexico and the United States[1]. During the first few weeks of surveillance, the virus spread worldwide to 30 countries (as of May 11) by human-to-human transmission, causing the World Health Organization to raise its pandemic alert to level 5 of 6. This virus has the potential to develop into the first influenza pandemic of the twenty-first century. Here we use evolutionary analysis to estimate the time-scale of the origins and the early development of the S-OIV epidemic. We show that it was derived from several viruses circulating in swine, and that the initial transmission to humans occurred several months before recognition of the outbreak. A phylogenetic estimate of the gaps in genetic surveillance indicates a long period of unsampled ancestry before the S-OIV outbreak, suggesting that the reassortment of swine lineages may have occurred years before emergence in humans, and that the multiple genetic ancestry of S-OIV is not indicative of an artificial origin. Furthermore, the unsampled history of the epidemic means that the nature and location of the genetically closest swine viruses reveal little about the immediate origin of the epidemic, despite the fact that we included a panel of closely related and previously unpublished swine influenza isolates. Our results highlight the need for systematic surveillance of influenza in swine, and provide evidence that the mixing of new genetic elements in swine can result in the emergence of viruses with pandemic potential in humans[2].

Initial genetic characterization of the S-OIV outbreak by the United States Centers for Disease Control suggested swine as its probable source, on the basis of sequence similarity to previously reported swine influenza isolates[1]. Classical swine H1N1 viruses have circulated in pigs in North America and other regions for at least 80 years[3]. In 1998, a new triple-reassortant H3N2 virus—comprising genes from classical swine H1N1, North American avian, and human H3N2 (A/Sydney/5/97-like) influenza—was reported as the cause of outbreaks in North American swine, with subsequent establishment in pig populations[4,5]. Co-circulation and mixing of the triple-reassortant H3N2 with established swine lineages subsequently generated further H1N1 and H1N2 reassortant swine viruses[6–8], which have caused sporadic human infections in the United States since 2005 (refs 6, 7). Consequently, human infection with H1N1 swine influenza has been a nationally notifiable disease in the United States since 2007 (ref. 9). In Europe, an avian H1N1 virus was introduced to pigs ('avian-like' swine H1N1) and first detected in Belgium in 1979 (ref. 10). This lineage became established and gradually replaced classical swine H1N1 viruses, and also reassorted in pigs with human H3N2 viruses (A/Port Chalmers/1/1973-like)[11]. It is noteworthy that, until now, there has been no evidence of Eurasian avian-like swine H1N1 circulating in North American pigs. In Asia, the classical swine influenza lineage circulates, in addition to other identified viruses, including human H3N2, Eurasian avian-like H1N1, and North American triple-reassortant H3N2 (refs 12, 13).

Using comprehensive phylogenetic analyses, we have estimated a temporal reconstruction of the complex reassortment history of the S-OIV outbreak, summarized in Fig. 1 (Methods). Our analyses showed that each segment of the S-OIV genome was nested within a well-established swine influenza lineage (that is, a lineage circulating primarily in swine for >10 years before the current outbreak). The most parsimonious interpretation of these results is therefore that the progenitor of the S-OIV epidemic originated in pigs. Some transmission of swine influenza has, however, been observed in secondary hosts in North America, for example, in turkeys[14]. Although the precise evolutionary pathway of the genesis of S-OIV is greatly hindered by the lack of surveillance data (see later), we can conclude that the polymerase genes, plus HA, NP and NS, emerged from a triple-reassortant virus circulating in North American swine. The source triple-reassortant itself comprised genes derived from avian (PB2 and PA), human H3N2 (PB1) and classical swine (HA, NP and NS) lineages. In contrast, the NA and M gene segments have their origin in the Eurasian avian-like swine H1N1 lineage. Phylogenetic analyses from the early days of the outbreak, on the basis of the first publicly available sequences, quickly established this multiple genetic origin (refs 8, 15, 16 and http://influenza.bio.ed.ac.uk).

Given that S-OIV contains genes of Eurasian origin, we included in our phylogenetic analyses 15 newly sequenced swine influenza viruses from Hong Kong, sampled in the course of a surveillance program conducted since the early 1990s. The viruses were a mixture of seven H1N1 and eight H1N2 subtypes, and viruses belonging to the classical, Eurasian avian-like, and triple-reassortant swine lineages were all present. Both Eurasian and triple-reassortant strains were isolated in Hong Kong in 2009. Extensive reassortment among these three virus lineages was also observed from the Hong Kong surveillance data (Supplementary Table 3), with reassortment between Eurasian avian-like and triple-reassortant swine lineages occurring as early as 2003 (for example, Sw/HK/78/2003).

Notably, for the PB1, HA and M genes, some of these newly generated sequences are more similar to the S-OIV epidemic than any previously reported isolates (Supplementary Fig. 2). Notably, seven out of eight genomic segments found in a single 2004 isolate (Sw/HK/915/04 (H1N2)) were located in a sister lineage to the current outbreak. Not only does this suggest that the precursors of S-OIV were swine viruses,

[1]State Key Laboratory of Emerging Infectious Diseases & Department of Microbiology, Li Ka Shing Faculty of Medicine, The University of Hong Kong, 21 Sassoon Road, Pokfulam, Hong Kong SAR, China. [2]Institute of Evolutionary Biology, University of Edinburgh, Ashworth Laboratories, King's Buildings, Edinburgh EH9 3JT, UK. [3]Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, Arizona 85705, USA. [4]Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, UK.

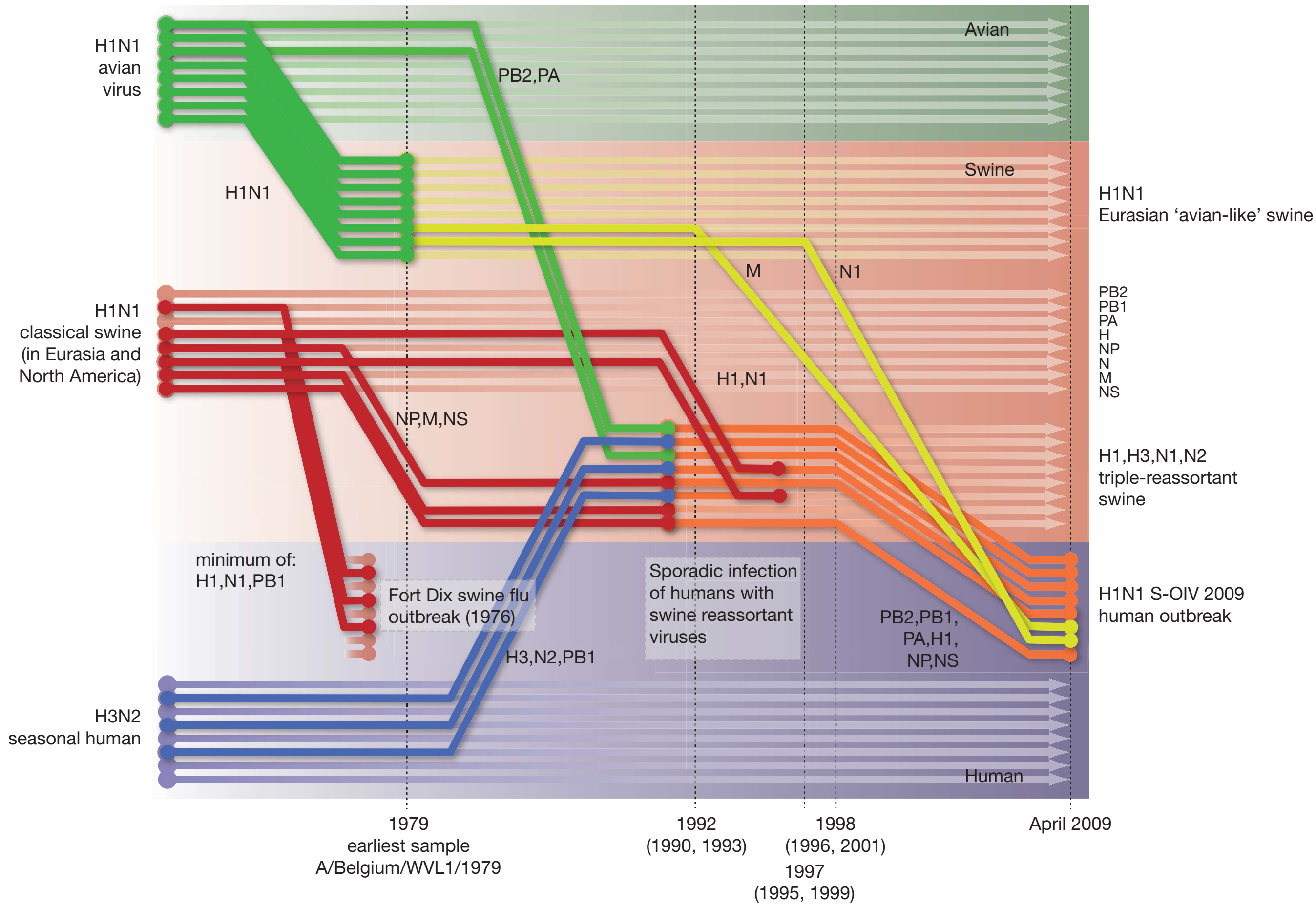


Figure 1 | Reconstruction of the sequence of reassortment events leading up to the emergence of S-OIV. Shaded boxes represent host species; avian (green), swine (red) and human (grey). Coloured lines represent interspecies-transmission pathways of influenza genes. The eight genomic segments are represented as parallel lines in descending order of size. Dates marked with dashed vertical lines on 'elbows' indicate the mean time of divergence of the S-OIV genes from corresponding virus lineages. Reassortment events not involved with the emergence of human disease are omitted. Fort Dix refers to the last major outbreak of S-OIV in humans. The first triple-reassortant swine viruses were detected in 1998, but to improve clarity the origin of this lineage is placed earlier.

but also that they were geographically widely distributed. Crucially, however, the observation of a sister relationship between the current outbreak virus and Sw/HK/915/04 cannot be interpreted as evidence for a Eurasian origin of the outbreak, owing to the long branch of the phylogeny leading to the 2009 human strains (Fig. 2 and Table 1). This branch must represent either an increased rate of evolution leading to the outbreak, or a long period during which the ancestors of the current epidemic went unsampled. To test these hypotheses, we regressed genetic divergence against sampling date for each gene, and found in favour of the latter: the evolutionary rate preceding the S-OIV epidemic is entirely typical for swine influenza (Supplementary Figs 2 and 3).

Therefore, to quantify the period of unsampled diversity, and to estimate the date of origin for the S-OIV outbreak, we performed a Bayesian molecular clock analysis for each gene (Methods). We also estimated the rate of evolution and time of the most recent common ancestor (TMRCA) of a set of genome sequences sampled from the S-OIV epidemic (between March and May 2009; isolates listed in Supplementary Table 4). We found that the common ancestor of the S-OIV outbreak and the closest related swine viruses existed between 9.2 and 17.2 years ago, depending on the genomic segment, hence the ancestors of the epidemic have been circulating undetected for about a decade. In contrast, the currently sampled S-OIV shared a common ancestor around January 2009 (no earlier than August 2008; Table 1). The long, unsampled history observed for every segment suggests that the reassortment of Eurasian and North American swine lineages may not have occurred recently, and it is possible that this single reassortant lineage has been cryptically circulating rather than two distinct lineages of swine flu. Thus, this genomic structure may have been circulating in pigs for several years before emergence in humans, and we urge caution in making inferences about human adaptation on the basis of the ancestry of the individual genes.

A search for amino acid residues in the S-OIV outbreak sequences that have been previously identified as phenotypic markers showed no evidence of virulence-associated variation or adaptations to human hosts[17–19], consistent with the outbreak being of swine origin and causing relatively mild symptoms. Full molecular characterization of the human swine H1N1 viruses is provided in Supplementary Information.

We did detect a difference in the viral molecular evolution in the outbreak clade when compared to that observed in related swine influenza sequences: all S-OIV genes showed a comparatively higher non-synonymous to synonymous (d_N/d_S) substitution rate ratio (Supplementary Tables 1 and 2). This d_N/d_S ratio rise could be due to the increased detection of mildly deleterious mutations resulting from intensive epidemic surveillance; such mutations would more typically be eliminated and escape detection[20]. Alternatively, these mutations could be adaptations to the new host species.

Because this d_N/d_S ratio rise may affect our estimate of the TMRCA of the S-OIV outbreak strains (which was estimated using long-term rates of swine influenza evolution), we compared the mean d_N/d_S values of outbreak versus non-outbreak data sets, thereby approximating the degree of excess of non-synonymous mutations in the outbreak sequences (Methods). Once the d_N/d_S ratio rise is corrected

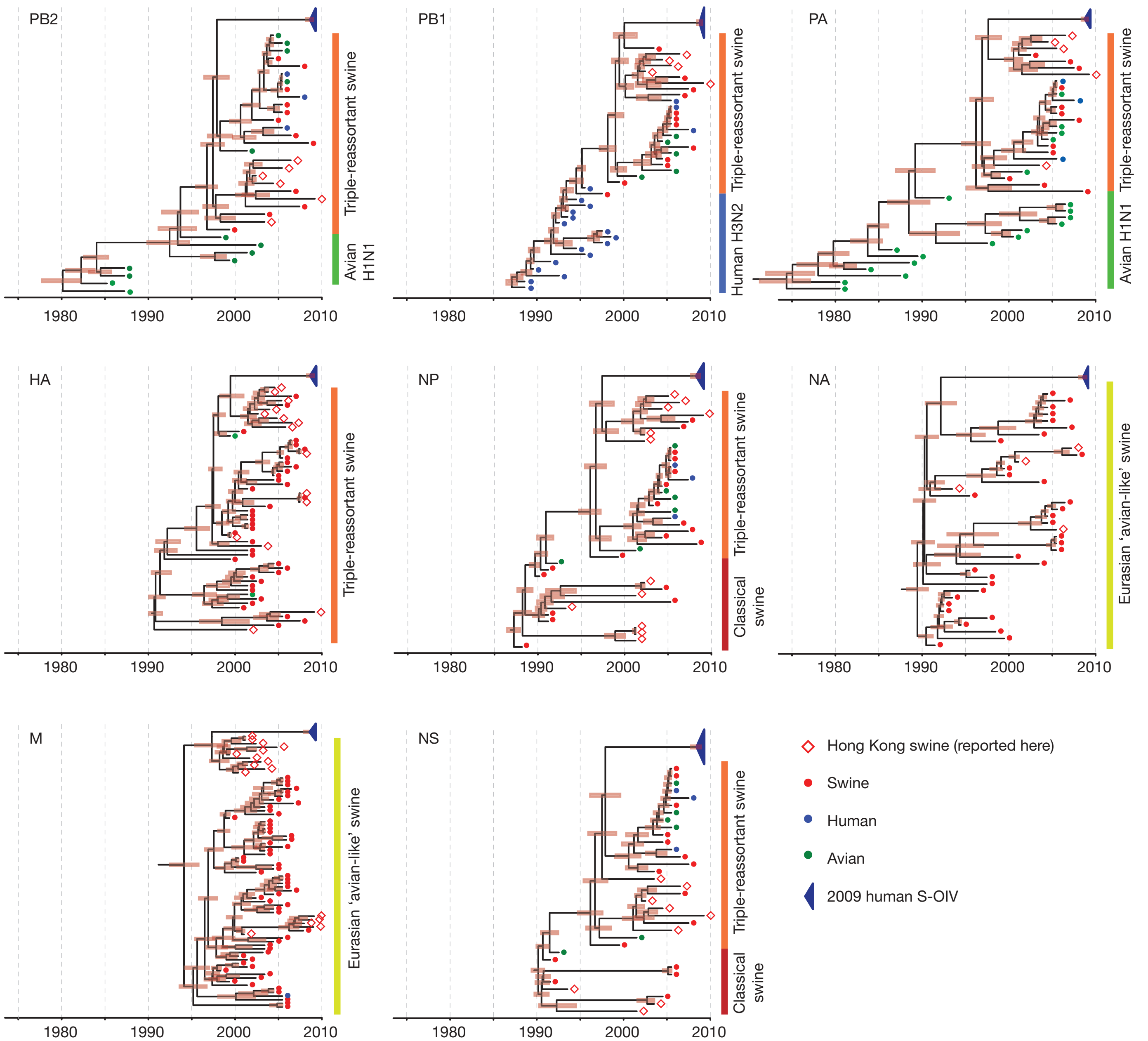


Figure 2 | Genetic relationships and timing of S-OIV for each genomic segment. Symbols represent sampled viruses on a timescale of when they were sampled and coloured by host species (pigs, red; humans, blue; birds, green). Internal nodes are reconstructed common ancestors with 95% credible intervals on their date given by the red bars. The S-OIV outbreak strains are represented by a blue triangle, with the apex representing the common ancestor of these.

for, the mean TMRCA of the S-OIV outbreak became 1 to 5 months more recent for each gene (Supplementary Tables 1 and 2). Furthermore, the adjusted TMRCA estimates are more uniform across genes, and are more similar to that obtained using internally calibrated S-OIV complete genomes (Table 1; a comparable estimate for the TMRCA of the HA gene only was recently reported[21]). Irrespective of whether the d_N/d_S ratio rise is due to increased detection of deleterious mutations or to increased adaptive evolution, its

Table 1 | Time of most recent common ancestors for the S-OIV outbreak

Gene	TMRCA of outbreak samples	Duration of unsampled diversity (years)	Mean evolutionary rate $\times 10^{-3}$ (subst. per site per year)
HA	28 Aug 2008 (1 Apr 2008, 2 Jan 2009)	9.80 (8.41, 11.02)	3.67 (3.41, 3.92)
MP	3 Aug 2008 (8 Dec 2007, 5 Feb 2009)	11.82 (10.17, 13.74)	2.55 (2.19, 2.93)
NA	8 Aug 2008 (23 Feb 2008, 26 Dec 2008)	17.15 (15.40, 18.88)	3.65 (3.22, 4.12)
NP	27 Mar 2008 (15 Sep 2007, 19 Sep 2008)	11.83 (10.53, 13.23)	2.59 (2.34, 2.84)
NS	21 May 2008 (30 Sep 2007, 27 Nov 2008)	11.47 (9.75, 13.21)	2.62 (2.32, 2.92)
PA	7 Oct 2008 (1 Jun 2008, 1 Feb 2009)	11.70 (10.25, 13.10)	2.45 (2.20, 2.69)
PB1	24 Oct 2008 (8 Jul 2008, 25 Jan 2009)	9.24 (7.59, 10.48)	2.34 (2.13, 2.53)
PB2	9 Sep 2008 (12 Apr 2008, 9 Jan 2009)	11.26 (9.93, 12.69)	2.60 (2.29, 2.92)
Genome*	21 Jan 2009 (3 Aug 2008, 13 Mar 2009)	N/A	3.66 (0.61, 6.58)

The values in parentheses represent the 95% credible intervals.
* This data set comprises complete or partial genomes of swine-origin influenza A (H1N1) virus outbreak isolates sampled predominantly in the United States between March and May 2009.

presence may be a general feature of intensively sampled emerging epidemics, and should be accounted for in the evolutionary analysis of such events.

Movement of live pigs between Eurasia and North America seems to have facilitated the mixing of diverse swine influenza viruses, leading to the multiple reassortment events associated with the genesis of the S-OIV strain. Domestic pigs have been described as a hypothetical 'mixing-vessel', mediating by reassortment the emergence of new influenza viruses with avian or avian-like genes into the human population, and triggering a pandemic associated with antigenic shift[2]. Previous research has suggested that occupational exposure to pigs increases the risk of swine influenza virus infection, and that swine workers should be considered in any surveillance programs[22].

The emergence of S-OIV provides further evidence of the role of domestic pigs in the ecosystem of influenza A. As reported recently, all three pandemics of the twentieth century seem to have been generated by a series of multiple reassortment events in swine or humans, and to have emerged over a period of years before pandemic recognition[23]. Our results show that the genesis of the S-OIV epidemic followed a similar evolutionary pathway: H1N1 viruses with human pandemic potential had been identified, transmission from swine to humans was known[5] and the disease had been made notifiable. Yet despite widespread influenza surveillance in humans, the lack of systematic swine surveillance allowed for the undetected persistence and evolution of this potentially pandemic strain for many years.

METHODS SUMMARY

We compared 15 newly sequenced Hong Kong swine influenza genomes and two genomes from the S-OIV outbreak with 796 genomes representing the spectrum of influenza A diversity (comprising 285 human, 100 swine and 411 avian isolates). Phylogenetic trees were constructed for each genomic segment independently (Supplementary Fig. 1). Next, for each genomic segment, viruses with known isolation dates that were genetically similar to the current outbreak were identified, and more detailed analysis using a Bayesian 'relaxed molecular clock' approach was performed[24], thereby estimating rates of viral evolution and dates of divergence (Fig. 2). Finally, a similar Bayesian molecular clock approach was applied to the 30 individual viruses isolated from the human outbreak since the end of March 2009 (Supplementary Table 4 and Supplementary Fig. 2). This analysis was performed assuming a model of exponential growth in the number of infections.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 24 May; accepted 4 June 2009.
Published online 11 June 2009.

1. Centers for Disease Control and Prevention. Swine influenza A (H1N1) infection in two children—Southern California, March–April 2009. *Morb. Mortal. Wkly Rep.* **58,** 400–402 (2009).
2. Shortridge, K. F., Webster, R. G., Butterfield, W. K. & Campbell, C. H. Persistence of Hong Kong influenza virus variants in pigs. *Science* **196,** 1454–1455 (1977).
3. Shope, R. E. & Lewis, P. Swine influenza: experimental transmission and pathology. *J. Exp. Med.* **54,** 349–359 (1931).
4. Brown, I. H., Harris, P. A., McCauley, J. W. & Alexander, D. J. Multiple genetic reassortment of avian and human influenza A viruses in European pigs, resulting in the emergence of an H1N2 virus of novel genotype. *J. Gen. Virol.* **79,** 2947–2955 (1998).
5. Webby, R. J. *et al.* Evolution of swine H3N2 influenza viruses in the United States. *J. Virol.* **74,** 8243–8251 (2000).
6. Newman, A. P. *et al.* Human case of swine influenza A (H1N1) triple reassortant virus infection, Wisconsin. *Emerg. Infect. Dis.* **14,** 1470–1472 (2008).
7. Shinde, V. *et al.* Triple-reassortant swine influenza A (H1) in humans in the United States, 2005–2009. *N. Engl. J. Med.* doi:10.1056/NEJMoa0903812 (in the press).
8. Novel Swine-Origin Influenza A (H1N1) Virus Investigation Team. Emergence of a novel swine-origin influenza A (H1N1) virus in humans. *N. Engl. J. Med.* doi:10.1056/NEJMoa0903810 (in the press).
9. Centers for Disease Control and Prevention. Novel influenza A virus infections—2007 case definition. <http://www.cdc.gov/ncphi/disss/nndss/casedef/novel_influenzaA.htm> (24 May 2009).
10. Pensaert, M., Ottis, K., Vanderputte, J., Kaplan, M. M. & Buchmann, P. A. Evidence for the natural transmission of influenza A virus from wild ducks to swine and its potential for man. *Bull. World Health Organ.* **59,** 75–78 (1981).
11. Brown, I. H. The epidemiology and evolution of influenza viruses in pigs. *Vet. Microbiol.* **74,** 29–46 (2000).
12. Peiris, J. S. M. *et al.* Cocirculation of avian H9N2 and contemporary "human" H3N2 influenza A viruses in pigs in southeastern China: potential for genetic reassortment? *J. Virol.* **75,** 9679–9686 (2001).
13. Jung, K. & Song, D. S. Evidence of the cocirculation of influenza H1N1, H1N2 and H3N2 viruses in the pig population of Korea. *Vet. Rec.* **161,** 104–105 (2007).
14. Choi, Y. K. *et al.* H3N2 influenza virus transmission from swine to turkeys, United States. *Emerg. Infect. Dis.* **10,** 2156–2160 (2004).
15. Trifonov, V., Khiabanian, H., Greenbaum, B. & Rabadan, R. The origin of the recent swine influenza A (H1N1) virus infecting humans. *Euro Surveill.* **14,** 19193 (2009).
16. Garten, R. J. *et al.* Antigenic and genetic characteristics of swine-origin 2009 A (H1N1) influenza viruses circulating in humans. *Science* doi:10.1126/science.1176225 (in the press).
17. Hatta, M., Gao, P., Halfmann, P. & Kawaoka, Y. Molecular basis of high virulence of Hong Kong H5N1 influenza A viruses. *Science* **7,** 1840–1842 (2001).
18. Le, Q. M., Sakai-Tagawa, Y., Ozawa, M., Ito, M. & Kawaoka, Y. Selection of H5N1 influenza virus PB2 during replication in humans. *J. Virol.* **83,** 5278–5281 (2009).
19. Obenauer, J. C. *et al.* Large-scale sequence analysis of avian influenza isolates. *Science* **311,** 1576–1580 (2006).
20. Pybus, O. G. *et al.* Phylogenetic estimation of deleterious mutation load in RNA viruses and its contribution to viral evolution. *Mol. Biol. Evol.* **24,** 845–852 (2007).
21. Fraser, C. *et al.* Pandemic potential of a strain of influenza A (H1N1): Early findings. *Science* doi:10.1126/science.1176062 (in the press).
22. Myers, K. P. *et al.* Are swine workers in the United States at increased risk of infection with zoonotic influenza virus? *Clin. Infect. Dis.* **42,** 14–20 (2006).
23. Smith, G. J. D. *et al.* Dating the emergence of pandemic influenza viruses. *Proc. Natl Acad. Sci. USA.* (in the press).
24. Drummond, A. J. & Rambaut, A. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* **7,** 214 (2007).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank E. C. Holmes for comments and encouragement. We acknowledge support from The Royal Society of London (A.R. and O.G.P.), the National Institute of Allergy and Infectious Diseases (NIAID) (G.J.D.S. and M.W.), the Biotechnology and Biological Sciences Research Council (BBSRC) (S.J.L.), and the David and Lucile Packard Foundation (M.W.). A.R. works as a part of the Interdisciplinary Centre for Human and Avian Influenza Research (ICHAIR). This study was supported by the National Institutes of Health (NIAID contract HHSN266200700005C) and the Area of Excellence Scheme of the University Grants Committee (grant AoE/M-12/06) of the Hong Kong SAR Government.

Author Contributions J.B., S.J.L., O.G.P., A.R., G.J.D.S., D.V. and M.W. conceived the study, performed analyses, co-wrote the paper, and all contributed equally to this work. J.S.M.P. co-wrote the paper, Y.G. conceived the study and co-wrote the paper, S.B. and J.R. performed analyses, S.K.M. conducted surveillance, and C.L.C. conducted sequencing. All authors commented on and edited the paper.

Author Information Newly reported sequences have been deposited at GenBank under accession numbers GQ229259–GQ229378. Reprints and permissions information is available at www.nature.com/reprints. This paper is distributed under the terms of the Creative Commons Attribution-Non-Commercial-Share Alike licence, and is freely available to all readers at www.nature.com/nature. Correspondence and requests for materials should be addressed to A.R. (a.rambaut@ed.ac.uk) or Y.G. (yguan@hku.hk).

METHODS

Sequence selection for phylogenetic analysis. We downloaded 3,986 complete influenza genomes of any subtype and sampling year (2,490 human, 185 swine and 1,311 avian) from the NCBI Influenza Virus Resource[25] on 29 April 2009. Each sequence set was given a unique ID of the form (ID number)_(Subtype)_(Host)_(isolate name), in which the isolate name is in lower case.

To reduce the number of very similar sequences, we listed all isolates in which the coding region in segment 1 (PB2) was at least one nucleotide different from the others. This left 1,759 human, 166 swine and 1,117 avian complete genome sets. Next we sampled the human, swine and avian sets, selecting one genome set per specific host (as defined in the isolate name, for example, chicken, duck), per specific location (for example, state or province), per year (although isolate name synonyms, for example, duck = dk, hongkong = hk were not accounted for). Two avian and four swine sequence sets were removed owing to bad sequences in one or more segments (for example, frameshifts), leaving 286 human (including S-OIVs), 100 swine and 411 avian sequences in the sampled subset. A further outbreak sequence set (A/Canada-ON/RV1527/2009), and the 15 new swine sequence sets were also added, making a total of 813 complete genome sets for analysis. For the more detailed, temporal analyses, all available S-OIV sequences were used.

The nucleotides in the coding regions of segments 1 (PB2), 2 (PB1), 3 (PA) and 5 (NP) were aligned using ClustalW[26] followed by manual alignment to codon position. The full nucleotide sequences of segments 7 (M1 and M2) and 8 (NS1 and NS2) were also aligned using ClustalW, and the sequences were edited such that all of the codons in first open reading frame (ORF) were followed by the remaining codons in the second ORF (that is, nucleotides were not repeated between the two ORFs). The HA and NA genes (segments 4 and 6) were aligned to codon positions using Muscle[27]. Further H1, H3, N1 and N2 only alignments were also performed.

New swine influenza sequences from Hong Kong. To evaluate the evolutionary history of swine/human influenza A H1N1 viruses, 15 viruses isolated from swine in Hong Kong during 1993 to 2009 were sequenced. Viral RNA was directly extracted from infected allantoic fluid or cell culture using QIAamp viral RNA minikit (Qiagen, Inc.). Complementary DNA was synthesized by reverse transcription reaction, and gene amplification by PCR was performed using specific primers for each gene segment. PCR products were purified with the QIAquick PCR purification kit (Qiagen Inc.) and sequenced by synthetic oligonucleotides. Reactions were performed using Big Dye-Terminator v3.1 Cycle Sequencing Reaction Kit on an ABI PRISM 3730 DNA Analyser (Applied Biosystems) following the manufacturer's instructions. All sequences were assembled and edited with Lasergene version 8.0 (DNASTAR). Full genome sequences of these viruses are available for download at GenBank under accession numbers GQ229259–GQ229378.

Molecular evolution and adaptation. We used the programs SLAC (Single-Likelihood Ancestor Counting)[28] and SNAP (Synonymous Non-synonymous Analysis Program)[29] to compare the mean ratio of non-synonymous changes per non-synonymous site to synonymous changes per synonymous site (d_N/d_S) of outbreak versus non-outbreak sequences. SLAC calculates inferred ancestral sequences for each internal node in a phylogeny using a codon model (and disallowing stop codons), and then counts the synonymous and non-synonymous mutations by comparing each codon to its immediate ancestor. SNAP counts the possible synonymous and non-synonymous codon changes across all pairs of sequences.

In brief, we calculated the effect of the excess of non-synonymous changes in the outbreak data as follows. Assume that S is the number of synonymous sites in a data set, N is the number of non-synonymous sites (typically $\sim$3.5S for these data), and ω is the d_N/d_S ratio. If the proportional contribution to the overall rate from synonymous sites is s, then the proportional contribution to the overall rate from non-synonymous sites is equal to $(N/S)(\omega)s$. N, S and ω are all readily estimated from the data. Assuming the same rate of synonymous substitution in both the outbreak and reference data sets, the relative rate expected in the outbreak sequences compared to the reference sequences is thus equal to

$$(s+(N/S)(\omega_{\text{outbreak}})s)/(s+(N/S)(\omega_{\text{reference}})s)$$

Phylogenetic analyses. Phylogenetic trees were inferred using the neighbour-joining distance method, with genetic distances calculated by maximum likelihood under the Hasegawa–Kishino–Yano (HKY) model with gamma-distributed rates among sites (HKY+Γ). Parameters of this model were estimated using maximum likelihood on an initial tree. Temporal phylogenies and rates of evolution were inferred using a relaxed molecular clock model that allows rates to vary among lineages within a Bayesian Markov chain Monte Carlo (MCMC) framework[24]. This was used to sample phylogenies and the dates of divergences between viruses from their joint posterior distribution, in which the sequences are constrained by their known date of sampling. A model comprising a codon-position-specific HKY+Γ substitution model was used. The limited sampling timespan of the S-OIV samples required a simpler model to avoid over-parameterization, so a single HKY+Γ model over all sites was used. For the analyses using Bayesian MCMC sampling, in all cases chain lengths of at least 50 million steps were used with a 10% 'burn-in' removed. Furthermore, at least two independent runs of each were performed and compared to ensure adequate sampling.

25. Bao, Y. *et al.* The influenza virus resource at the national center for biotechnology information. *J. Virol.* **82**, 596–601 (2008).
26. Thompson, J. D., Higgins, D. G. & Gibson, T. J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673–4680 (1994).
27. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797 (2004).
28. Kosakovsky Pond, S. L. & Frost, S. D. W. Not so different after all: a comparison of methods for detecting amino acid sites under selection. *Mol. Biol. Evol.* **22**, 1208–1222 (2005).
29. Korber, B. *HIV Signature and Sequence Variation Analysis. Computational Analysis of HIV Molecular Sequences* (eds Rodrigo, A. G. & Learn, G. H.) Ch. 4, 55–72 (Kluwer Academic Publishers, 2000).

LETTERS

Down's syndrome suppression of tumour growth and the role of the calcineurin inhibitor DSCR1

Kwan-Hyuck Baek[1]*, Alexander Zaslavsky[1]*, Ryan C. Lynch[1]*, Carmella Britt[1], Yoshiaki Okada[2], Richard J. Siarey[3], M. William Lensch[4], In-Hyun Park[4], Sam S. Yoon[5], Takashi Minami[6], Julie R. Korenberg[7], Judah Folkman[1], George Q. Daley[4], William C. Aird[2], Zygmunt Galdzicki[3] & Sandra Ryeom[1]

The incidence of many cancer types is significantly reduced in individuals with Down's syndrome[1–4], and it is thought that this broad cancer protection is conferred by the increased expression of one or more of the 231 supernumerary genes on the extra copy of chromosome 21. One such gene is Down's syndrome candidate region-1 (*DSCR1*, also known as *RCAN1*), which encodes a protein that suppresses vascular endothelial growth factor (VEGF)-mediated angiogenic signalling by the calcineurin pathway[5–10]. Here we show that DSCR1 is increased in Down's syndrome tissues and in a mouse model of Down's syndrome. Furthermore, we show that the modest increase in expression afforded by a single extra transgenic copy of *Dscr1* is sufficient to confer significant suppression of tumour growth in mice, and that such resistance is a consequence of a deficit in tumour angiogenesis arising from suppression of the calcineurin pathway. We also provide evidence that attenuation of calcineurin activity by DSCR1, together with another chromosome 21 gene *Dyrk1a*, may be sufficient to markedly diminish angiogenesis. These data provide a mechanism for the reduced cancer incidence in Down's syndrome and identify the calcineurin signalling pathway, and its regulators DSCR1 and DYRK1A, as potential therapeutic targets in cancers arising in all individuals.

Down's syndrome is the most common genetic cause of mental retardation in humans, occurring in 1 out of 700 live births. Epidemiological studies suggest that although individuals with Down's syndrome have an increased risk of leukaemia, they have a considerably reduced incidence of most solid tumours[1–4]. In the largest study to date involving 17,800 Down's syndrome individuals, the mortality from cancers was $<$10% of expected[4]. Such data indicate that one or more of the 231 trisomic genes on chromosome 21 is responsible for protecting these individuals against cancer. Of note, Down's syndrome individuals also have a reduced incidence of other angiogenesis-related diseases, such as diabetic retinopathy[11] and atherosclerosis[12], suggesting that cancer protection in the Down's syndrome population may be due, in part, to angiogenesis suppression.

The *DSCR1* gene lies on chromosome 21 and encodes a negative regulator of VEGF-calcineurin signalling in the endothelium[5–10]. Previous studies have demonstrated that gross overexpression of DSCR1 in endothelial cells blocks VEGF-mediated angiogenic responses *in vitro*[7–10]. We examined DSCR1 protein expression in Down's syndrome human fetal tissues and observed a consistent 1.8-fold increase in DSCR1 levels in tissues from Down's syndrome embryos, as well as increased expression of DSCR1 fetal isoforms[5], compared with those from age-matched control fetuses (Fig. 1a). Substantial synteny exists in gene identity and order between much of human chromosome 21 and the mouse[13]. The Ts65Dn mouse model of Down's syndrome is trisomic for 104 of the 231 genes on human chromosome 21, including DSCR1 (ref. 14) (Supplementary Fig. 1a). We probed tissues from Ts65Dn mice to ascertain whether DSCR1 was also upregulated in the Ts65Dn mouse and found a 1.7-fold increase in DSCR1 protein expression compared with diploid littermates (Supplementary Fig. 1b). Thus, DSCR1 expression in Ts65Dn mice is increased in an analogous fashion to DSCR1 expression in Down's syndrome fetal tissues.

We proposed that the inhibition of tumorigenesis in the Down's syndrome population may be partially due to the suppression of tumour angiogenesis, thus predicting that the inhibitory effect of *DSCR1* trisomy on tumour growth occurs within the host tumour microenvironment. To establish whether the Ts65Dn Down's syndrome mouse, like Down's syndrome individuals, exhibits generalized protection from cancers, we assayed the growth of two transplantable tumour models: Lewis lung carcinoma and B16F10 melanoma cells. We observed considerable growth suppression of both Lewis lung and B16F10 tumour cells in Ts65Dn mice relative to littermate controls (Fig. 1b), correlating with a significant decrease in microvessel density (Fig. 1c). Endothelial cells isolated from Ts65Dn mice demonstrated upregulation of *Dscr1* mRNA in contrast to diploid littermates (Supplementary Fig. 1c) and were notably less responsive to VEGF-mediated proliferation *in vitro* (Supplementary Fig. 1d), further implicating an angiogenic defect in these mice. Thus, trisomy for orthologues of half the genes on human chromosome 21 was sufficient to slow ectopic tumour growth.

To validate that the compromised angiogenesis we observed in murine models of Down's syndrome extended to human cells carrying trisomy 21, we compared microvessel density in teratomas derived from Down's syndrome induced pluripotent stem (iPS) cells versus those from cytogenetically normal iPS cells from a healthy volunteer[15]. These iPS cell lines were inoculated intramuscularly into immunodeficient mice (*Rag2*$^{-/-}$ *Il2rg*$^{-/-}$), and angiogenesis was assessed in the resulting tumours using a human-specific antibody to the endothelial marker CD31 (ref. 16). Microvessel density was significantly reduced in teratomas derived from Down's syndrome iPS cells compared to iPS cells from the normal control (Fig. 1d and Supplementary Fig. 1e).

[1]Vascular Biology Program, Department of Surgery, Children's Hospital Boston, Massachusetts 02115, USA. [2]Center for Vascular Biology Research, Division of Molecular and Vascular Medicine, Beth Israel Deaconess Medical Center, Boston, Massachusetts 02115, USA. [3]Department of Anatomy, Physiology and Genetic, Neuroscience, Molecular and Cellular Biology Program, School of Medicine, Uniformed Services University of the Health Services, Bethesda, Maryland 20814, USA. [4]Department of Medicine, Division of Pediatric Hematology Oncology, Children's Hospital Boston, and Dana-Farber Cancer Institute, Boston, Massachusetts 02115, USA. [5]Division of Surgical Oncology, Department of Surgery, Massachusetts General Hospital, Boston, Maryland 02114, USA. [6]Research Center for Advanced Science and Technology, University of Tokyo, Tokyo, 153-8904, Japan. [7]Department of Pediatrics and The Brain Institute, The University of Utah, Salt Lake City, Utah 84108, USA.
*These authors contributed equally to this work.

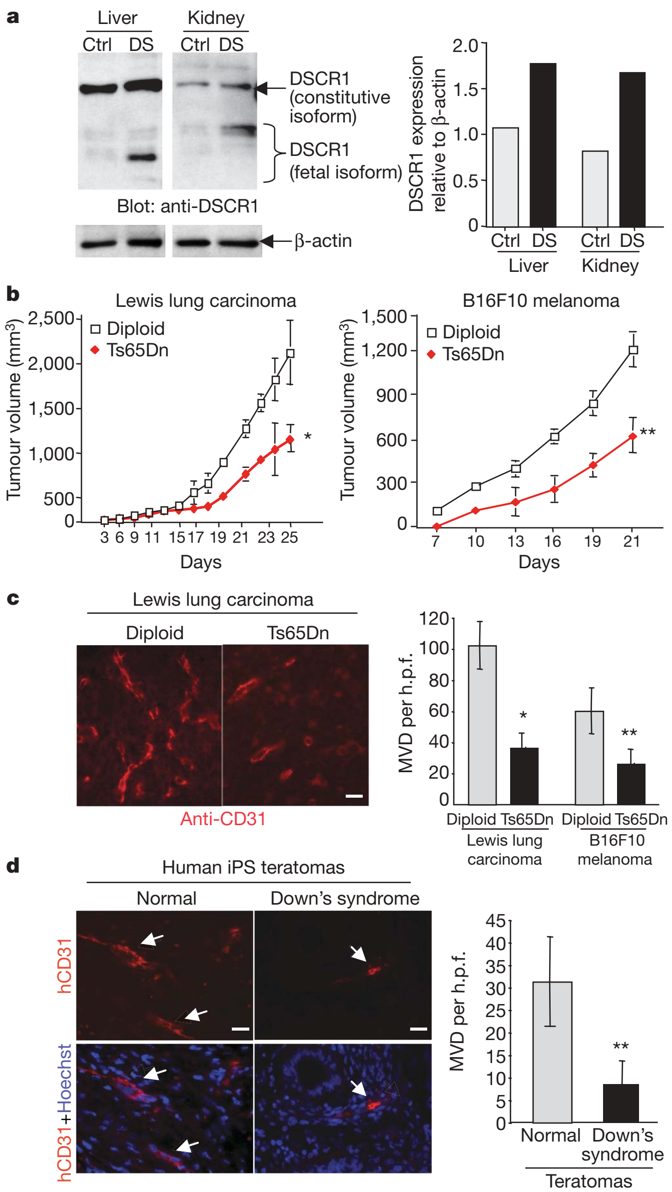


Figure 1 | *Dscr1* expression is upregulated and tumour angiogenesis is suppressed in Down's syndrome tissues and models. **a**, Increased DSCR1 expression in human fetal Down's syndrome (DS) tissues versus age-matched control (ctrl) tissues, relative to β-actin. **b**, Tumour growth is suppressed in the Ts65Dn Down's syndrome mouse model. $n = 10$–12; $*P < 0.03$, $**P < 0.01$. **c**, Microvessel density (MVD) per high-powered field (h.p.f.) of tumours is quantified by anti-CD31 immunofluorescence. Scale bar, 20 μM. $*P < 0.02$, $**P < 0.01$. **d**, Angiogenesis in tumours from iPS cells isolated from Down's syndrome or cytogenetically normal cells was quantified by human-specific anti-CD31 immunofluorescence. Arrows denote hCD31-positive vessels. $n = 3$–6; $**P < 0.01$. All values are mean ± s.e.m.

Given that Ts65Dn mice are trisomic for 104 genes, we next determined whether three copies of *Dscr1* alone would be sufficient to suppress tumorigenesis. We generated a *Dscr1* transgenic mouse targeting a Myc-tagged *Dscr1* complementary DNA driven by its native promoter to the *Hprt1* locus (Supplementary Fig. 2a, b). Expression of the targeted third copy of *Dscr1* was verified by western blot analysis on isolated endothelial cells (Supplementary Fig. 2c), and quantitative PCR analysis demonstrated a 2.4-fold increase in *Dscr1* messenger RNA relative to littermate controls (Supplementary Fig. 2d). To ensure that one extra copy of *Dscr1* was sufficient to restrain VEGF-calcineurin signalling, we examined NF-ATc1 subcellular localization as a measure of calcineurin activation (Fig. 2a).

After VEGF treatment, endothelial cells isolated from *Dscr1* transgenic mice exhibited predominantly cytoplasmic NF-ATc1 localization, whereas wild-type endothelial cells showed the expected nuclear localization (Fig. 2a). *Dscr1* transgenic endothelial cells also exhibited decreased sensitivity to VEGF relative to wild-type endothelial cells, as assessed by VEGF-induced proliferation (Fig. 2b).

Substantial growth inhibition of Lewis lung carcinoma and B16F10 melanoma cells was observed in *Dscr1* transgenic mice (Fig. 2c), with a corresponding decrease in microvessel density (Fig. 2d), as compared with tumours from wild-type littermates. Quantification of endothelial cells in tumours isolated from *Dscr1* transgenic and wild-type mice by flow cytometry showed a significant decrease in $CD31^{+}CD45^{-}$ cells in tumours from *Dscr1* transgenic mice (Supplementary Fig. 3a). Furthermore, many of the CD31-reactive microvessels in tumours isolated from *Dscr1* transgenic mice lacked functional lumens as evidenced by the absence of co-staining with circulating fluorescein isothiocyanate (FITC)–lectin and CD31-positivity (Supplementary Fig. 3b). Immunostaining with CD31 and Myc antibodies confirmed specific expression of the DSCR1 transgene in the tumour endothelium of *Dscr1* transgenic animals (Fig. 2e). Thus, a single extra copy of *Dscr1* seems to be sufficient to blunt host angiogenic responses and suppress tumour angiogenesis and tumour growth. A subset of tumours initiate growth by co-opting existing host vessels[17] but the progressive growth and elaboration of new vessels accompanying tumour expansion requires neo-angiogenesis. To ascertain whether excess DSCR1 suppresses initial vessel co-option or subsequent angiogenesis, *Dscr1* transgenic and wild-type mice were inoculated with reduced numbers of Lewis lung carcinoma cells to generate slowly growing tumours. DSCR1 increase inhibited the extended growth phase of transplanted tumours and not just their initial expansion (Supplementary Fig. 3c), supporting the notion that excess DSCR1 suppresses tumour angiogenesis.

To confirm whether tumour protection in the Ts65Dn mouse was specifically due to *Dscr1* trisomy, we crossed the Ts65Dn mice to $Dscr1^{+/-}$ mice[18] generating Ts65Dn/$Dscr1^{+/-}$ animals with two copies of *Dscr1* but maintaining trisomy for the other 103 genes (Fig. 3a). After validating both segmental trisomy and *Dscr1* status (Supplementary Fig. 4a, b), we compared flank tumour growth in diploid mice to that in Ts65Dn mice with two or three copies of *Dscr1*. Reduction to two *Dscr1* copies in the Ts65Dn/$Dscr1^{+/-}$ mice significantly abrogated the tumour protection observed in the Ts65Dn parental mice (Fig. 3b). Loss of the protective effect against tumour growth was mirrored by a corresponding increase in microvessel density in tumours from Ts65Dn/$Dscr1^{+/-}$ mice relative to their Ts65Dn littermates (Fig. 3c), confirming the pivotal role of *Dscr1* in tumour suppression in the Ts65Dn Down's syndrome mouse model. Together, our data provide strong support for the notion that one extra copy of *Dscr1* is necessary for maximal suppression of tumour growth via inhibition of tumour angiogenesis.

Increased dosage of *Dscr1* might suppress angiogenesis by down-regulating expression of calcineurin-NFAT-dependent targets in endothelial cells. Cyclooxygenase 2 (COX-2, also known as PTGS2) has been identified as a calcineurin-dependent gene and an important mediator of the angiogenic response to VEGF[19]. Quantitative PCR and western blot analysis of COX-2 expression in *Dscr1* transgenic endothelial cells demonstrated a substantial decrease in COX-2 levels relative to control littermates (Fig. 4a, b), suggesting that a modest increase in DSCR1 expression suppresses the expression of COX-2 and probably other VEGF-responsive targets. Although *Dscr1* trisomy has an important role in preventing tumour angiogenesis in our Down's syndrome mouse model, there are clearly other chromosome 21 genes that contribute to the tumour suppressive effects observed in the Ts65Dn mouse. Because increased dosage of *Dscr1* attenuates VEGF-calcineurin-NFAT signalling, we examined the role of *Dyrk1a*, another chromosome 21 gene known to regulate NFAT signalling and contribute to other Down's syndrome phenotypes[20,21]. After confirmation of *Dyrk1a* expression in endothelial cells

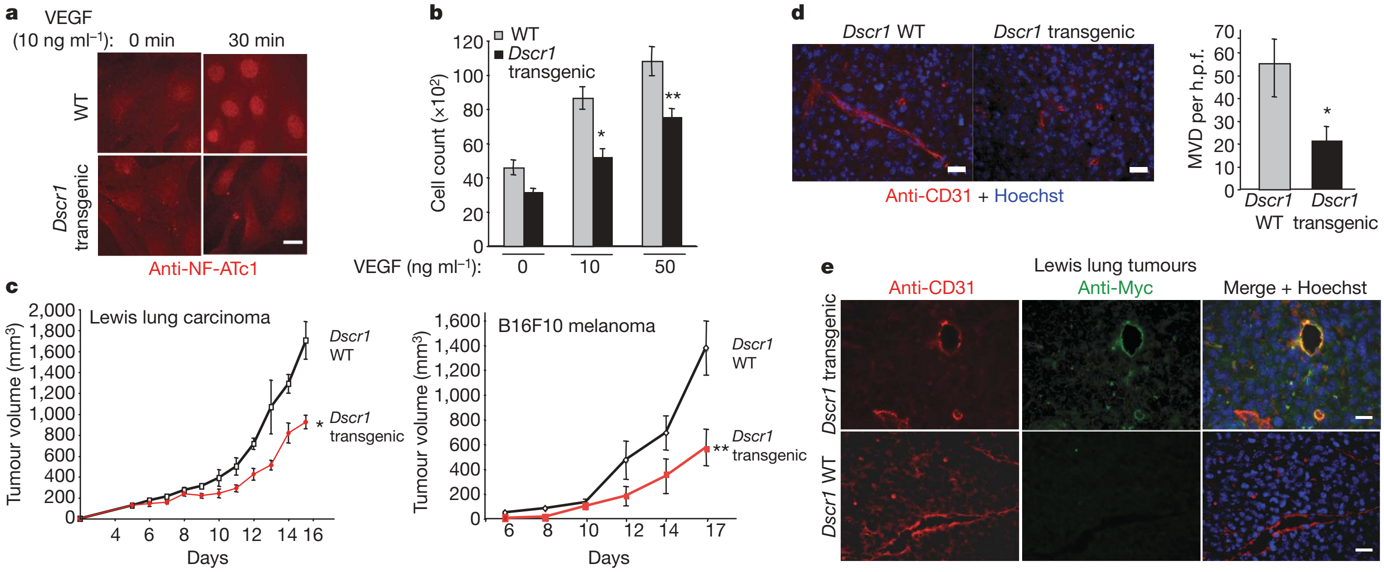


Figure 2 | Targeted *Dscr1* transgenic mice with three copies of *Dscr1* show inhibition of tumour growth. a, NF-ATc1 nuclear import after VEGF treatment is suppressed in endothelial cells from *Dscr1* transgenic mice. Scale bar, 5 μM. **b**, VEGF-induced proliferation of endothelial cells from *Dscr1* transgenic mice is significantly inhibited. WT, wild type. $*P < 0.02$, $**P < 0.03$. **c**, Tumour growth is suppressed in *Dscr1* transgenic mice. $n = 8$–12 mice per group; $*P < 0.05$, $**P < 0.04$. **d**, Microvessel density per high-powered field of tumours is quantified by anti-CD31 immunofluorescence. Scale bars, 20 μM. Values are mean ± s.e.m. $*P < 0.01$. **e**, Tumours from *Dscr1* transgenic and wild-type mice are co-immunostained with anti-CD31 to detect endothelial cells and anti-Myc to detect the *Dscr1* transgene. Scale bars, 20 μM.

(Supplementary Fig. 4c), we examined the effects of *Dyrk1a* upregulation with *Dscr1* trisomy after overexpression of *Dyrk1a* into *Dscr1* transgenic endothelial cells (Fig. 4c). VEGF-mediated endothelial proliferation was markedly inhibited as a consequence of increased expression of both DYRK1A and DSCR1 (Fig. 4d), indicating that upregulation of *Dyrk1a* may contribute to the remaining tumour suppression observed in the Ts65Dn/*Dscr1*$^{+/-}$ mouse with two copies of *Dscr1*. Collectively, our data demonstrate that the modest excess of DSCR1 afforded by a single extra copy impedes VEGF-calcineurin signalling in endothelial cells. The consequent suppression of tumour angiogenesis and, ultimately, tumour growth provides a mechanistic basis for the remarkable protection from solid tumours observed in Down's syndrome individuals. Of note, host *Dscr1* copy number influences the growth and vascularization of allografted tumour cell lines, confirming the cell intrinsic tumour suppressive effect of *Dscr1* dosage by modulating angiogenesis within the host tumour microenvironment.

Because human chromosome 21 contains over 200 genes, it would be surprising if *Dscr1* was the only chromosome 21 gene implicated in tumour suppression in Down's syndrome individuals. Indeed our data suggest that increased dosage of *Dyrk1a* seems to act together with *Dscr1* to suppress tumour angiogenesis by further attenuating

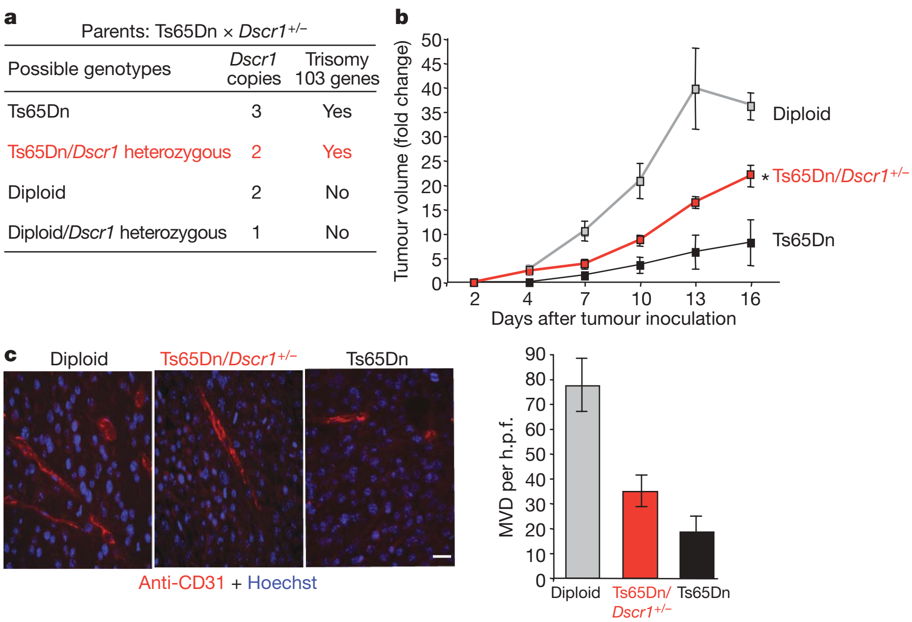


Figure 3 | Trisomic expression of *Dscr1* is necessary for significant suppression of tumour growth in the Ts65Dn Down's syndrome mouse model. a, Possible genotypes arising from mating Ts65Dn and *Dscr1*$^{+/-}$ mice. **b**, Suppression of B16F10 tumour growth in Ts65Dn Down's syndrome mice is relieved after loss of the third copy of *Dscr1* (Ts65Dn/*Dscr1*$^{+/-}$). Values are mean ± s.e.m., $n = 4$–8 per group, $*P < 0.01$. **c**, Microvessel density per high-powered field is quantified by anti-CD31 immunofluorescence of tumours collected from the indicated mice at comparable volumes (200–400 mm^3). Scale bar, 20 μM. Values are mean ± s.e.m.

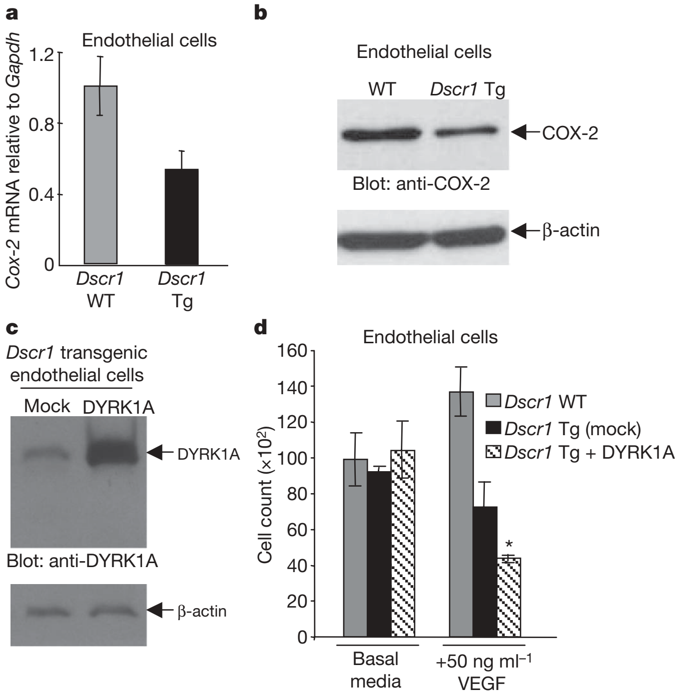


Figure 4 | Calcineurin suppression by DSCR1 and DYRK1A attenuates endothelial cell activation. a, **b**, *Cox-2* mRNA (**a**) and protein expression (**b**) were quantified by qPCR and western blot analysis in endothelial cells isolated from *Dscr1* wild-type (WT) and transgenic (Tg) mice. **c**, Western blot analysis of endothelial cells isolated from *Dscr1* transgenic mice and probed for DYRK1A after either mock infection or retroviral infection with *Dyrk1a*. **d**, Endothelial cells isolated from *Dscr1* transgenic mice with DYRK1A overexpression (+DYRK1A) demonstrate an even greater suppression of VEGF-mediated proliferation as compared to wild-type and DSCR1 overexpression alone (mock). Values are mean ± s.e.m., $*P < 0.01$.

VEGF-calcineurin-NFAT signalling in endothelial cells. By disrupting the balance of NFAT phosphorylation, DYRK1A blocks transactivation of NFAT-dependent target genes[20,21]. Other trisomy 21 genes may inhibit tumour development via cancer cell autonomous mechanisms. A recent study using the Ts65Dn mouse on the APC^{min} mice background demonstrated that three copies of the *Ets2* proto-oncogene reduced the incidence of spontaneous intestinal tumours[22]. However, examination of a spontaneous tumour model prevents the distinction of *Ets2* dosage effects on tumour initiation versus progression. In contrast, our work specifically investigates suppression of tumour progression as a consequence of *Dscr1* trisomy. Further studies are necessary to determine the magnitude and synergy of these genes in suppressing tumour growth in both the Ts65Dn mouse and the Down's syndrome population.

Our studies implicate DSCR1-dependent inhibition of the VEGF-calcineurin-NFAT pathway in endothelial cells as an important component of the reduced cancer incidence in Down's syndrome individuals. Hence, it is reasonable to speculate that cyclosporin A and FK506, immunosuppressive drugs that specifically inhibit calcineurin[23], would also suppress tumour angiogenesis. Surprisingly, several clinical studies indicate that a significant increase in cancer incidence is a serious complication of transplant recipients receiving long-term immunosuppressive therapy[24]. The mechanism behind this increased rate of cancer is not yet understood, however, such studies point to an important distinction between the calcineurin-inhibitory action of DSCR1 and those of cyclosporin A and FK506.

The cancer protection observed in Down's syndrome individuals is remarkable given the heterogeneous mechanisms used among tumours. Although Down's syndrome individuals have less exposure to environmental and other factors that contribute to tumour incidence, the implication remains that one or more of the trisomic genes on chromosome 21 exert an anti-neoplastic effect, presumably by modulating some fundamental aspect of tumour initiation and/or progression. Microscopic avascular tumours represent the earliest stages of human neoplasias and are commonly observed in many organs after autopsy[25]. These data suggest that progression into macroscopic tumours, not tumour initiation, may be rate-limiting in human cancers. Such progression is critically dependent on interactions between the incipient tumour and its microenvironment—most notably, activation of the 'angiogenic switch'. The pivotal role played by DSCR1 in tumour angiogenesis makes it a compelling candidate for cancer protection in the Down's syndrome population. Our data implicate VEGF-calcineurin-NFAT as a critical signalling axis in endothelial cells and suggests that maximal suppression of calcineurin requires increased expression of both *Dscr1* and *Dyrk1a*. Further studies will be required to explore whether angiogenesis in general may also be impaired in the Ts65Dn mouse and in Down's syndrome individuals. Finally, our studies in the Ts65Dn mouse and transgenic tri-allelic *Dscr1* mouse demonstrate the *in vivo* relevance of a modest excess of *Dscr1* arising from trisomy, which is sufficient to negatively regulate tumour growth by dampening VEGF-calcineurin signalling. By analogy, we conclude that the increased expression of *Dscr1* in individuals with trisomy 21 is probably a significant contributor to the decreased cancer incidence in this population. It is, perhaps, inspiring that the Down's syndrome population provides us with new insight into mechanisms that regulate cancer growth and, by so doing, identifies potential targets for tumour prevention and therapy.

METHODS SUMMARY

Western blot analysis. Fourteen-week-old human fetal kidney and liver tissue were isolated from control and Down's syndrome tissue, lysed in RIPA buffer and quantified for protein concentration. Tissue lysates were separated by SDS–PAGE, probed with anti-DSCR1 monoclonal antibody[18] and detected by chemiluminescence. Blots were stripped and re-probed with β-actin. DSCR1 levels were quantified by densitometric analysis in the linear range and compared to β-actin. Ts65Dn mice, diploid littermates and *Dscr1*-null mice were euthanized between 8 to 12 weeks of age, their brains were dissected and probed as described above. Human microvascular endothelial cells were lysed and probed with anti-DYRK1A monoclonal antibody.

Tumour models. Six–ten-week-old mice were inoculated with $1–5 \times 10^5$ Lewis lung carcinoma or B16F10 melanoma cells in HBSS into the subcutaneous flank region as previously described[6]. Human iPS cell-derived teratomas were generated as previously described[15].

Immunohistochemistry. Mice were killed and tumours were collected and fixed in neutral-buffered formalin for paraffin embedding. Sections were deparaffinized and epitopes unmasked as previously described[6]. Sections were immunostained with rat anti-CD31 monoclonal antibody or anti-Myc polyclonal antibody overnight at room temperature, followed by incubation with goat anti-rat Alexa 594 and goat anti-rabbit Alexa 488, stained with Hoechst and analysed using AxioVision 4.0 software (Carl Zeiss Vision).

Endothelial cell isolation, immunofluorescence and proliferation. Four-week-old *Dscr1* transgenic or littermate control mice were killed, lungs were removed, and endothelial cells were isolated as previously described[6]. For immunofluorescence, endothelial cells were plated on gelatine-coated coverslips, fixed with paraformaldehyde, blocked, permeabilized and incubated with anti-NFATc1 monoclonal antibody followed by goat anti-mouse Alexa 594. For proliferation assays, endothelial cells were plated in triplicate in tissue culture wells coated with 0.2% gelatin as previously described[6], and treated with the indicated concentrations of VEGF.

Statistical analysis. Data are shown as mean ± s.e.m. or s.d. as denoted. *P* values were calculated using Student's *t*-test.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 15 January; accepted 16 April 2009.
Published online 20 May 2009.

1. Hasle, H., Clemmensen, I. H. & Mikkelsen, M. Risks of leukaemia and solid tumors in individuals with Down's syndrome. *Lancet* **355,** 165–169 (2000).
2. Satge, D. *et al.* A lack of neuroblastoma in Down syndrome: a study from 11 European countries. *Cancer Res.* **58,** 448–452 (1998).

3. Patja, K., Pukkala, E., Sund, R., Iivanainen, M. & Kaski, M. Cancer incidence of persons with Down syndrome in Finland: a population-based study. *Int. J. Cancer* **118,** 1769–1772 (2006).
4. Yang, Q., Rasmussen, S. A. & Friedman, J. M. Mortality associated with Down's syndrome in the USA from 1983 to 1997: a population-based study. *Lancet* **359,** 1019–1025 (2002).
5. Fuentes, J. J., Pritchard, M. A. & Estivill, X. Genomic organization, alternative splicing, and expression patterns of the *DSCR1* (Down syndrome candidate region 1) gene. *Genomics* **44,** 358–361 (1997).
6. Ryeom, S. *et al.* Targeted deletion of the calcineurin inhibitor DSCR1 suppresses tumor growth. *Cancer Cell* **13,** 420–431 (2008).
7. Hesser, B. A. *et al.* Down syndrome critical region protein 1 (DSCR1), a novel VEGF target gene that regulates expression of inflammatory markers on activated endothelial cells. *Blood* **104,** 149–158 (2004).
8. Iizuka, M., Abe, M., Shiiba, K., Sasaki, I. & Sato, Y. Down syndrome candidate region 1, a downstream target of VEGF, participates in endothelial cell migration and angiogenesis. *J. Vasc. Res.* **41,** 334–344 (2004).
9. Minami, T. *et al.* Vascular endothelial growth factor- and thrombin-induced termination factor, Down syndrome critical region-1, attenuates endothelial cell proliferation and angiogenesis. *J. Biol. Chem.* **279,** 50537–50554 (2004).
10. Yao, Y. G. & Duh, E. J. VEGF selectively induces Down syndrome critical region 1 gene expression in endothelial cells: a mechanism for feedback regulation of angiogenesis? *Biochem. Biophys. Res. Commun.* **321,** 648–656 (2004).
11. Fulcher, T. *et al.* Diabetic retinopathy in Down's syndrome. *Br. J. Ophthalmol.* **82,** 407–409 (1998).
12. Murdoch, J. C., Rodger, J. C., Rao, S. S., Fletcher, C. D. & Dunnigan, M. G. Down's syndrome: an atheroma-free model? *BMJ* **2,** 226–228 (1977).
13. Olson, L. E., Richtsmeier, J. T., Leszl, J. & Reeves, R. H. A chromosome 21 critical region does not cause specific Down syndrome phenotypes. *Science* **306,** 687–690 (2004).
14. Reeves, R. H. *et al.* A mouse model for Down syndrome exhibits learning and behaviour deficits. *Nature Genet.* **11,** 177–184 (1995).
15. Park, I. H. *et al.* Disease-specific induced pluripotent stem cells. *Cell* **134,** 877–886 (2008).
16. Melero-Martin, J. M. *et al.* Engineering robust and functional vascular networks *in vivo* with human adult and cord blood-derived progenitor cells. *Circ. Res.* **103,** 194–202 (2008).
17. Holash, J. *et al.* Vessel cooption, regression, and growth in tumors mediated by angiopoietins and VEGF. *Science* **284,** 1994–1998 (1999).
18. Ryeom, S., Greenwald, R. J., Sharpe, A. H. & McKeon, F. The threshold pattern of calcineurin-dependent gene expression is altered by loss of the endogenous inhibitor calcipressin. *Nature Immunol.* **4,** 874 (2003).
19. Hernandez, G. L. *et al.* Selective inhibition of vascular endothelial growth factor-mediated angiogenesis by cyclosporin A: roles of the nuclear factor of activated T cells and cyclooxygenase 2. *J. Exp. Med.* **193,** 607–620 (2001).
20. Arron, J. R. *et al.* NFAT dysregulation by increased dosage of *DSCR1* and *DYRK1A* on chromosome 21. *Nature* **441,** 595–600 (2006).
21. Gwack, Y. *et al.* A genome-wide *Drosophila* RNAi screen identifies DYRK-family kinases as regulators of NFAT. *Nature* **441,** 646–650 (2006).
22. Sussan, T. E., Yang, A., Li, F., Ostrowski, M. C. & Reeves, R. H. Trisomy represses Apc^{Min}-mediated tumours in mouse models of Down's syndrome. *Nature* **451,** 73–75 (2008).
23. Liu, J. *et al.* Calcineurin is a common target of cyclophilin-cyclosporin A and FKBP-FK506 complexes. *Cell* **66,** 807–815 (1991).
24. Dantal, J. & Soulillou, J. P. Immunosuppressive drugs and the risk of cancer after organ transplantation. *N. Engl. J. Med.* **352,** 1371–1373 (2005).
25. Black, W. C. & Welch, H. G. Advances in diagnostic imaging and overestimations of disease prevalence and the benefits of therapy. *N. Engl. J. Med.* **328,** 1237–1243 (1993).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank G. Evan, C. Kim, K. Cichowski and J. Italiano for critical discussions and advice. This work was supported by the Howard Hughes Medical Institute (G.Q.D.), Harvard Stem Cell Institute (G.Q.D.), the National Institute of Health Director's Pioneer Award (G.Q.D.), NHLBI (W.C.A.), Jerome Lejeune Foundation (Z.G.), USUHS (Z.G.), the Smith Family Medical Foundation (S.R.), the Garrett B. Smith Foundation (S.R.) and Annie's Fun Foundation (S.R.).

Author Contributions K.-H.B., A.Z., R.C.L. and C.B. performed the experiments, analysed the data and generated the figures. I.-H.P. and M.W.L. performed experiments. Y.O., R.J.S., T.M., W.C.A. and Z.G. generated the transgenic mouse models in this manuscript and analysed data. G.Q.D. designed experiments and analysed data. S.S.Y. and J.R.K. contributed to pilot experiments. J.F. and S.R. generated the hypotheses. S.R. designed and performed experiments, analysed the data and wrote the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to S.R. (sandra.ryeom@childrens.harvard.edu).

METHODS

Generation of Ts65Dn mice and Ts65Dn/*Dscr1* heterozygous mice. Two-month-old diploid and Ts65Dn mice age 2–4 months old were obtained by breeding Ts65Dn females with C57BL/6C3H F_1 males, or were acquired from Jackson Laboratories. Mice were karyotyped with chromosomal spreads from the blood[26], and metaphase chromosomal spreads were prepared and karyotypes evaluated. All protocols were approved by the Children's Hospital or USUHS Institutional Animal Care and Use Committee. Experiments were performed without previous knowledge of genotypes.

The Ts65Dn/*Dscr1*$^{+/-}$ mice that were diploid for the *Dscr1* gene and trisomic for the rest of the Ts65Dn segment of chromosome 16 were generated as follows: males heterozygous for the *Dscr1* gene on a C57Bl/6 background were bred with Ts65Dn females. Litters were genotyped by fluorescence *in situ* hybridization (FISH) and PCR screening using tail DNA as previously described[6,27]. In brief, nuclei of blood leukocytes stained blue by 4,6-diamidino-2-phenylindole (DAPI) show red spots that reveal the presence of chromosomes 16 and Ts65Dn segment, detected by FISH with a probe produced by labelling BAC 480C6 (Research Genetics; containing a fragment of Ts65Dn segment) with biotin-dUTP or digoxigenin-dUTP (Boehringer Mannheim) in a nick-translation reaction.

***Dscr1* transgenic mice.** A 593-base-pair (bp) cDNA fragment encoding the inducible murine *Dscr1* isoform (*Dscr1.Ex4*) was subcloned with a 5′ 6-*myc* epitope into an *Hprt* targeting vector with the *Dscr1.Ex4* native 2 kb promoter[28]. This allowed the generation of single-copy, site-specific recombination at the *Hprt* locus which reduced the risk of nonspecific phenotypes due to site of insertion[29]. This construct was targeted into the *Hprt* locus by electroporation into *Hprt*-deficient embryonic stem cells (BK4 cells). Two independent recombinant embryonic stem cell clones screened by PCR and Southern blot were injected into blastocysts, with both lines demonstrating germline transmission. *Dscr1* transgenic mice were backcrossed onto a C57Bl/6 background for six generations.

Western blot analysis. Fourteen-week-old human fetal kidney and liver tissue were isolated from control and Down's syndrome tissue obtained from therapeutic miscarriages according to an approved Institutional Review Board (IRB) protocol at UCLA. Tissues were lysed in RIPA buffer and protein quantified by the BioRad DC Protein Assay. Five micrograms of tissues were separated by SDS–PAGE, probed with anti-DSCR1 (clone 14B4) monoclonal antibody[18] and detected via chemiluminescence (ECL, Amersham). Blots were stripped and re-probed with β-actin. Levels of DSCR1 were quantified by densitometric analysis in the linear range and compared to β-actin. Ts65Dn mice, diploid littermates and *Dscr1*-null mice were killed between 8 to 12 weeks of age, their brains were dissected and probed as described above.

Human microvascular endothelial cells (Cambrex) were cultured as previously described[6], lysed, separated by SDS–PAGE and probed with anti-DYRK1A polyclonal antibody (Abcam). Primary mouse endothelial cells from the indicated genotypes were lysed, separated by SDS–PAGE and probed with anti-Myc (9E10) or anti-COX2 monoclonal antibody (Alexis).

Immunohistochemistry and microvessel density. Mice were killed and tumours were collected and fixed in neutral-buffered formalin for paraffin embedding. Paraffin embedded sections were deparaffinized and epitopes unmasked as previously described[6]. Sections were immunostained with rat anti-CD31 monoclonal antibody (1:50; Pharmingen) overnight at room temperature followed by incubation for 2 h with goat anti-rat Alexa 594 conjugated secondary antibody (1:1,000, Molecular Probes), and analysed using AxioVision 4.0 software. Microvessel density was quantified after CD31 immunostaining of five sections per tumour, five mice per cohort and quantified as previously described[6].

Quantification of microvessel density in tumours by flow cytometry. Tumour-bearing mice were anesthetized with avertin and the vasculature perfused for 3 min at 120 mm Hg pressure with saline from an 18-gauge cannula inserted into the aorta via an incision in the left ventricle. Tumours were dissected, minced with scissors, and then digested in PBS supplemented with 2 mg ml^{-1} of collagenase/dispase (Roche) for 45 min at 37 °C. Dissociated cells were resuspended with PBS containing 10% FBS and 10^6 cells were stained with anti-mouse CD31–FITC-conjugated antibody plus anti-mouse CD45-PerCP-conjugated antibody (BD Pharmingen) for 15 min on ice. Cells were washed in PBS, examined in a FACScan flow cytometer, and analysed using CellQuest software (BD Biosciences).

FITC–lectin perfusion. Mice were anesthetized with isofluorane. FITC-labelled Lycopersicon esculentum lectin (100 μg in 100 μl of 0.9% NaCl; Vector Laboratories) was injected into the femoral vein and allowed to circulate for 3 min before perfusion of fixative. The chest was opened and the vasculature perfused for 3 min at a pressure of 120 mm Hg with fixative (4% paraformaldehyde in PBS) from an 18-gauge cannula inserted into the aorta via an incision in the left ventricle. The right atrium was incised to create a route for the fixative to escape. After removal, tissues were processed for immunohistochemistry.

Immunofluorescence. One million endothelial cells per ml were plated onto gelatin-coated glass coverslips overnight. Cells were fixed with 4% paraformaldehyde for 10 min followed by blocking in 3% milk in TBS-T for 45 min. Anti-VEGF-R2 monoclonal antibody (1:500; Cell Signaling) or anti-mouse NF-ATc1 (1:200) or isotype-matched control antibodies was added for 1 h. Cells were washed with PBS-T before the addition of rabbit anti-mouse-Alexa594 (1:500; Molecular Probes) for 30 min protected from light. Nuclei were stained with 1% Hoechst dye for 1 min. Cells were washed with PBS-T, mounted and imaged on a fluorescence microscope.

Quantification of *Dscr1* and *Cox-2* mRNA in endothelial cells. Total RNA was isolated from 5×10^5 endothelial cells using the RNeasy kit (Qiagen). cDNA was generated from 5 μg total RNA using SuperScript III First-Strand Synthesis system (Invitrogen) and subjected to real-time PCR with PerfeCTa SYBR Green SuperMix (Quanta Biosciences) using the RNA Engine Opticon Monitor 2 System (MJ Research Inc.). *Dscr1* was amplified with forward primer 5′-AGCTCCCTGATTGCTTGTGT-3′ and reverse primer 5′-AGGAACTCGGT CTTGTGCAG-3′. *Cox-2* was amplified with forward primer 5′-CCCCCAC AGTCAAAGACACT-3′ and reverse primer 5′-AGTTGCTCATCACCCCA CTC-3′. *Gapdh* was amplified with forward primer 5′-TGAAGGTCGGT GTGAACGGATTTGG-3′ and reverse primer 5′-CATGTAGGCCATGAGGT CCACCAC-3′.

Retroviral expression of DYRK1A. *Dyrk1a* cDNA was obtained from Open Biosystems, excised with EcoR1 and cloned into pBABE. DYRK1A retrovirus was generated by transfecting GPG293 cells as previously described[30]. Virus was used to infect 2×10^5 human microvascular endothelial cells or primary mouse endothelial cells by serial incubations with DYRK1A virus.

26. Harashima, C. *et al.* Elevated expression of the G-protein-activated inwardly rectifying potassium channel 2 (GIRK2) in cerebellar unipolar brush cells of a Down syndrome mouse model. *Cell. Mol. Neurobiol.* **26**, 719–734 (2006).
27. Shi, Y. P., Huang, T. T., Carlson, E. J. & Epstein, C. J. The mapping of transgenes by fluorescence in situ hybridization on G-banded mouse chromosomes. *Mamm. Genome* **5**, 337–341 (1994).
28. Yang, J. *et al.* Independent signals control expression of the calcineurin inhibitory proteins MCIP1 and MCIP2 in striated muscles. *Circ. Res.* **87**, E61–E68 (2000).
29. Bronson, S. K. *et al.* Single-copy transgenic mice with chosen-site integration. *Proc. Natl Acad. Sci. USA* **93**, 9067–9072 (1996).
30. Giuriato, S. *et al.* Sustained regression of tumors upon MYC inactivation requires p53 or thrombospondin-1 to reverse the angiogenic switch. *Proc. Natl Acad. Sci. USA* **103**, 16266–16271 (2006).

LETTERS

Biomechanical forces promote embryonic haematopoiesis

Luigi Adamo[1]*, Olaia Naveiras[2]*, Pamela L. Wenzel[2], Shannon McKinney-Freeman[2], Peter J. Mack[1], Jorge Gracia-Sancho[1], Astrid Suchy-Dicey[1], Momoko Yoshimoto[3], M. William Lensch[2], Mervin C. Yoder[3], Guillermo García-Cardeña[1]* & George Q. Daley[2]*

Biomechanical forces are emerging as critical regulators of embryogenesis, particularly in the developing cardiovascular system[1,2]. After initiation of the heartbeat in vertebrates, cells lining the ventral aspect of the dorsal aorta, the placental vessels, and the umbilical and vitelline arteries initiate expression of the transcription factor *Runx1* (refs 3–5), a master regulator of haematopoiesis, and give rise to haematopoietic cells[4]. It remains unknown whether the biomechanical forces imposed on the vascular wall at this developmental stage act as a determinant of haematopoietic potential[6]. Here, using mouse embryonic stem cells differentiated *in vitro*, we show that fluid shear stress increases the expression of *Runx1* in CD41$^+$c-Kit$^+$ haematopoietic progenitor cells[7], concomitantly augmenting their haematopoietic colony-forming potential. Moreover, we find that shear stress increases haematopoietic colony-forming potential and expression of haematopoietic markers in the para-aortic splanchnopleura/aorta–gonads–mesonephros of mouse embryos and that abrogation of nitric oxide, a mediator of shear-stress-induced signalling[8], compromises haematopoietic potential *in vitro* and *in vivo*. Collectively, these data reveal a critical role for biomechanical forces in haematopoietic development.

In the mouse, the first haemogenic areas appear in the yolk sac starting at day 7.5 of development (E7.5)[9]. After the establishment of circulation and the onset of vascular flow at day 8.5, additional haemogenic sites appear between day 9 and 10.5 as Runx1$^+$ regions within the developing vasculature[10]. Given the developmental and anatomical relationship of haematopoietic precursors with the mechano-responsive vascular endothelium[11,12], together with the temporal correlation between the establishment of circulation and the appearance and expansion of vascular haemogenic sites, we proposed that biomechanical forces might act to promote haematopoiesis.

To explore this hypothesis, we differentiated mouse embryonic stem (ES) cells as embryoid bodies, a system that recapitulates the early stages of embryonic haematopoietic development. In embryoid bodies, cells first commit to mesoderm and then produce Flk1$^+$ cells containing the earliest embryonic haematopoietic precursors[13]. Embryoid bodies were cultured until the appearance of Flk1$^+$ mesoderm on day 3.25 of differentiation[13,14], disaggregated and plated on flat gelatinized surfaces. To define the type of biomechanical stimulation to apply to these cultured cells, we focused on the haemodynamic environment present in the aorta–gonad–mesonephros (AGM) region, the best characterized haemogenic site at the onset of circulation[15,16]. The pulsatile characteristics of flow in the aorta generate a complex interplay of distinct types of biomechanical forces including circumferential stress, hydrodynamic pressure and shear stress. Of these forces, we focused on fluid shear stress, the frictional force generated by viscous flow acting along cells lining blood vessels, because it has been shown to exert profound effects on the structure and function of vascular endothelial cells[12]. We therefore chose to stimulate cultured cells using a wall shear stress (WSS) comparable to that acting along the dorsal aorta throughout the cardiac cycle at E10.5. To estimate the time-average of this value, we used previously published ultrasound biomicroscopy (UBM) data for the time-average velocity over one cardiac cycle (velocity time integral, 4.74 mm, divided by the cycle length, 350 ms, and the earliest available values for dorsal aorta diameter, 0.33 mm (refs 10, 17)). Calculating the circular Poiseuille flow solution for wall shear stress with these values yielded a dorsal aorta WSS at E10.5 of approximately 5 dyn cm^{-2} (Supplementary Fig. 1a). This approach assumed that the UBM velocity data represented mean blood flow and that Poiseuille flow assumptions are valid within the E10.5 dorsal aorta[18] (see Methods).

Embryoid-body-derived cells were exposed to this biomechanical stimulus (WSS) or cultured under static conditions for 48 h (Fig. 1a and Supplementary Fig. 1b). Exposure of embryoid-body-derived cells to shear stress caused an increase in cells positive for CD31 (PECAM1), a marker of endothelial[19] and haematopoietic lineages (Supplementary Fig. 1c). CD41$^+$c-Kit$^+$ haematopoietic precursors[7] were sorted by fluorescence-activated cell sorting (FACS). Gene expression analysis within this compartment demonstrated a strong shear-stress-mediated upregulation of the transcription factors *Runx1* (also called *Cbfa2*; 4.6-fold) and *Myb* (2.8-fold), the prototypical markers of haemogenic sites[4], and of *Klf2* (4.1-fold), a gene previously shown to be a driver of erythropoiesis[20] and to be mechano-activated in endothelial cells[21] (Fig. 1b). *Runx1* upregulation was observed in the unsorted cell population and this effect was specific to the WSS values estimated for the AGM region in the early embryo. Indeed, when the embryoid-body-derived cells were exposed to a shear stress of different magnitude or to a shear stress waveform characteristic of the human aorta the expression of Runx1 was not increased (Supplementary Fig. 1d).

Next, we interrogated the functional significance of these gene expression changes using haematopoietic colony-forming assays. As a positive control, we used the iCdx4 ES cell line, which carries a doxycycline-inducible *Cdx4* transgene that enhances differentiation to haematopoietic precursors[22]. Shear stress increased the frequency of haematopoietic colony-forming units (c.f.u.) when compared to static (no shear stress) conditions, with a magnitude comparable to that obtained by induction of *Cdx4* with doxycycline (Fig. 1c). Notably, the *Cdx4* transgene was not activated in cells

[1]Center for Excellence in Vascular Biology, Departments of Pathology Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts 02115, USA. [2]Stem Cell Transplantation Program, Division of Pediatric Hematology/Oncology, Children's Hospital Boston and Dana Farber Cancer Institute; Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School; Division of Hematology, Brigham and Women's Hospital; Harvard Stem Cell Institute; Manton Center for Orphan Disease Research; Howard Hughes Medical Institute, Boston, Massachusetts 02115, USA. [3]Department of Pediatrics, Indiana University School of Medicine, Indianapolis, Indiana 46202, USA.
*These authors contributed equally to this work.

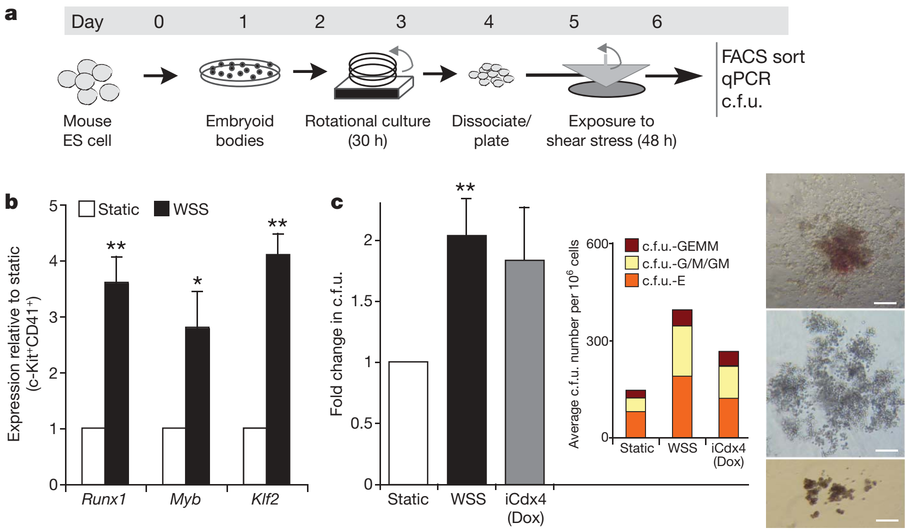


Figure 1 | Shear stress induces haematopoietic commitment from ES-derived cells. **a**, Experimental protocol used to induce haematopoietic differentiation from ES-derived cells in the presence of wall shear stress (WSS). ES cells are differentiated with the embryoid body method for 3.25 days, disaggregated and plated on gelatinized surfaces. Cell monolayers were exposed to shear stress and then collected on day 6 for further analysis **b**, Real-time Taqman PCR-based gene expression analysis in FACS-sorted CD41$^+$c-Kit$^+$ embryoid-body-derived haematopoietic precursors. Exposure to WSS induces upregulation of the haematopoietic markers *Runx1* ($P = 0.01$), *Myb* ($P = 0.03$) and *Klf2* ($P = 0.001$), $n = 3$. **c**, Methylcellulose haematopoietic c.f.u. assay. WSS increases the frequency of haematopoietic progenitors in complete M3434 methylcellulose. $n = 4$, analysis of variance (ANOVA) $P = 0.01$. Dox, doxycycline. Inset shows average distribution of haematopoietic colony types. c.f.u.-GEMM, c.f.u. granulocyte-erythroid-myeloid-megakaryocytes; c.f.u.-G/M/GM, c.f.u. granulocytes/macrophages/granulocyte-macrophages; c.f.u.-E, c.f.u. erythroid. Bar graphs represent average ± s.e.m. Pictures show representative colonies: c.f.u.-GEMM (top), c.f.u.-G/M/GM (middle), c.f.u.-E (bottom). Scale bar, 200 μm. $*P < 0.05$, $**P < 0.01$.

exposed to shear stress (Supplementary Fig. 1e). We observed similar results when we analysed Flk1$^+$ cells enriched by magnetic cell sorting, indicating that the shear stress response occurs in the Flk1$^+$ mesoderm (Supplementary Fig. 1f). Importantly, the small number of detached cells in static or shear conditions was very similar (<0.5%) and the haematopoietic potential in this non-adherent fraction was higher in the cells exposed to WSS (Supplementary Fig. 1g). These observations demonstrate that biomechanical forces enhance embryonic stem-cell-derived haematopoiesis.

To gain mechanistic insights into this process, we investigated the role of nitric oxide (NO), a well characterized signalling pathway strongly regulated by shear stress[8], and a known modulator of haematopoiesis[23]. Inhibition of NO production via the nitric oxide synthase inhibitor nitro-L-arginine methyl ester (L-NAME)[24] resulted in a significant reduction of the shear-stress-induced enrichment in c.f.u. as compared with cells treated with the inactive stereoisomer D-NAME (Fig. 2a). Notably, L-NAME did not affect shear-stress-induced *Runx1* upregulation in CD41$^+$c-Kit+ cells, indicating that shear-stress-mediated Runx1 upregulation is upstream of NO production (Fig. 2b). L-NAME and D-NAME had no effect on c.f.u. formation or *Runx1* expression in cells grown under static conditions (Supplementary Fig. 2a, b).

We next sought to determine the role of NO production on intraembryonic haematopoiesis *in vivo*. We administered L-NAME or D-NAME to pregnant mice for 48 h, starting from the time of establishment of circulation (E8.5), and assessed the number of c.f.u. in the AGM regions of E10.5 embryos. As shown in Fig. 2c, systemic inhibition of NO production led to a marked decrease in the number of c.f.u. per AGM. In this *in vivo* setting, we cannot distinguish between the direct effect of NO inhibition on haematopoietic precursors and the alterations of embryonic haemodynamics that result from changes in vascular tone triggered by the systemic inhibition of NO synthases[24]. However, our *in vitro* data document that NO production is required for the shear-stress-mediated stimulation of haematopoietic progenitors,

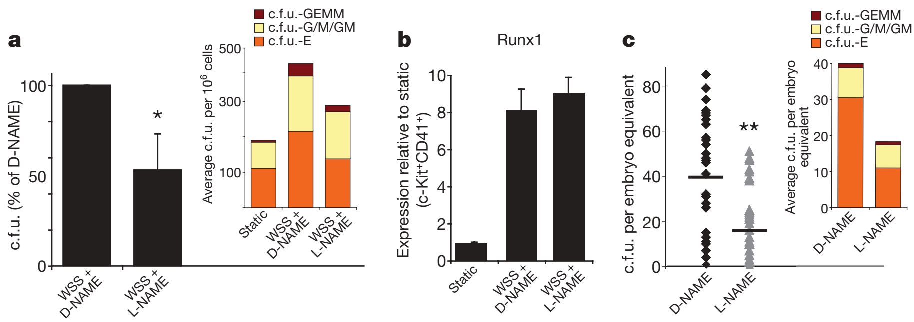


Figure 2 | Nitric oxide production regulates the expansion of haematopoietic progenitors. **a**, Methylcellulose haematopoietic c.f.u. assay. Pharmacological inhibition of nitric oxide synthesis with L-NAME reduces by 50% the WSS-mediated increase in haematopoietic c.f.u.; $n = 3$, $P = 0.04$. Inset shows the average distribution of colony types. **b**, L-NAME does not affect WSS-mediated Runx1 upregulation; $n = 3$. **c**, *In vivo* c.f.u. assay. Exposure of developing embryos to L-NAME from E8.5 to E10.5 leads to a reduction in haematopoietic progenitors in the AGM region. D-NAME, $n = 30$; L-NAME, $n = 32$; $P = 0.0008$. Inset shows average distribution of colonies per AGM. Bar graphs represent average ± s.e.m. $*P < 0.05$, $**P < 0.005$.

thereby establishing that the NO pathway is an important mediator of the effect of shear stress in haematopoiesis.

To examine the effect of fluid shear stress on the haematopoietic forming potential of embryonic haematopoietic sites, we established two-dimensional primary cultures of para-aortic splanchnopleura (PSp, the precursor of the AGM in E9–E9.5 embryos) or AGM-derived cells. When PSp regions from E9.5 murine embryos were disaggregated, plated and exposed to shear stress, we observed an increase in c.f.u. when compared to static controls (Fig. 3a). Likewise, when we exposed cells isolated from the AGM regions of E10.5 murine embryos to shear stress, we observed enhanced expression of Runx1 and Klf2, but for unclear reasons not Myb, in $CD41^{+}$ haematopoietic progenitors enriched by FACS (Fig. 3b)[25]. Shear stress also induced upregulation of $CD31^{+}$ cells (Fig. 3c). We then examined the effect of shear stress on specific haematopoietic lineages by flow cytometry, and documented an increase in cells positive for the B lymphocyte marker B220, and Ter119, a marker of erythroblasts (Fig. 3d and Supplementary Fig. 3a). Furthermore, the erythroblasts present in the shear-stress-treated cultures displayed morphological features consistent with later stages of erythroid maturation (that is, pycnotic erythroblasts and a haemoglobinized red pellet versus earlier polychromatic erythroblasts present in the static control; Fig. 3d and Supplementary Fig. 3b), suggesting accelerated erythroid differentiation in response to shear stress. This was associated with strong upregulation of *Klf2*, a known driver of erythropoiesis. Taken together, these observations show that shear stress increases both the prevalence of haematopoietic progenitors and the expression of haematopoietic markers in primary cultures of cells taken from the PSp/AGM, indicating that shear stress acts to enhance embryonic haematopoiesis.

To evaluate further the relevance of shear stress in the maturation of embryonic haemogenic sites *in vivo*, we assessed the effect of fluid shear stress in mouse embryos carrying a homozygous mutation in the Na^{+}/Ca^{2+} exchanger *Ncx1* (also called *Slc8a1*). These embryos fail to initiate the heartbeat and therefore lack circulation[26]. $Ncx1^{-/-}$ embryos are deficient in haematopoietic c.f.u. in the PSp, possibly due to a failure of redistribution from the yolk sac[27], and have a 6-fold reduction in $CD41^{+}$ haematopoietic progenitors in the placental vasculature[5]. The data presented here suggest that the haematopoietic deficiencies observed in $Ncx1^{-/-}$ embryos may be due, at least in part, to a lack of biomechanical stimulation at haemogenic vascular sites. To evaluate this possibility, we determined the levels of *Runx1* and *Klf2* expression in the PSp/AGM region of E9.25 $Ncx1^{-/-}$ embryos and wild-type littermates. As shown in Fig. 4a, both *Runx1* and *Klf2* were expressed in $Ncx1^{-/-}$ PSp/AGM region at significantly lower levels than in controls, whereas expression of the endothelial-specific marker VE-cadherin was similar in the two groups, suggesting that the content of vascular endothelial tissue was similar in both samples (data not shown). Importantly, at this developmental stage, *Runx1* is exclusively expressed in haemogenic sites within the embryo trunk[28,29]. We then examined if exposure to biomechanical stimulation per se could restore expression of Runx1 and haematopoietic colony-forming potential in $Ncx1^{-/-}$ embryo-derived cells. We isolated cells from the PSp/AGM region of $Ncx1^{-/-}$ embryos, and either maintained them under static conditions or exposed them to shear stress. As shown in Fig. 4, shear stress induced Runx1 expression (Fig. 4b) and c.f.u. activity (Fig. 4c) in cells from $Ncx1^{-/-}$ embryos. These data demonstrate that shear stress is capable of upregulating expression of the haematopoietic master regulator *Runx1* and promoting haematopoietic colony-forming activity in

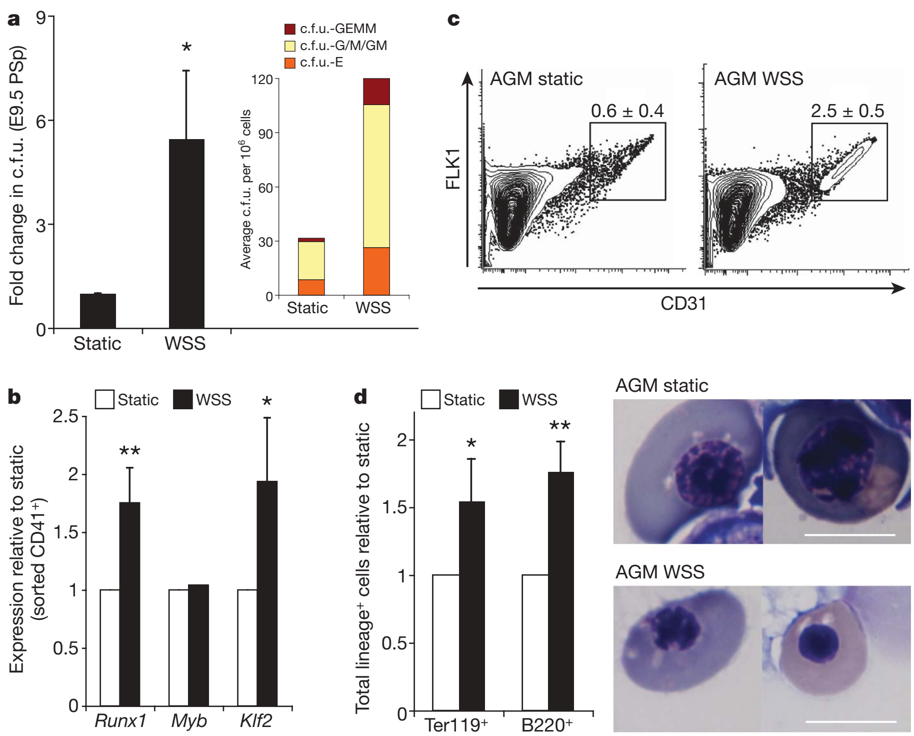


Figure 3 | Shear stress induces haematopoiesis in PSp/AGM-embryo-derived cells. a, Methylcellulose haematopoietic c.f.u. assay. WSS increases the frequency of haematopoietic progenitors in two-dimensional primary PSp cultures from E9.5 embryos; $n = 4$ ($P = 0.038$); bar graphs represent average $\pm$ s.e.m. Inset shows average distribution of haematopoietic colony types. **b**, WSS induces upregulation of the haematopoietic markers *Runx1* ($P = 0.01$) and *Klf2* ($P = 0.05$) in FACS-sorted AGM-derived $CD41^{+}$ haematopoietic progenitors as documented by real-time PCR; $n = 3$, bar graphs represent average $\pm$ s.e.m. **c**, FACS analysis. WSS induces an increase in $CD31^{+}$ cells in two-dimensional primary AGM cultures; $P = 0.005$, $n = 3$, average $\pm$ s.d. **d**, WSS modulates the differentiation of AGM-derived haematopoietic progenitors as shown by an increase in absolute number of cells positive for the erythroid marker Ter119 ($P = 0.02$) and for the lymphoid marker B220 ($P = 0.01$); $n = 3$, bar graphs represent average $\pm$ s.e.m. Shear stress induces maturation of erythroid precursors as documented by cell morphology in cytospins, which show pycnotic erythroblasts in WSS-treated samples, and polychromatic erythroblasts in static cultures. Scale bar, 10 μm. $^{*}P < 0.05$, $^{**}P < 0.01$.

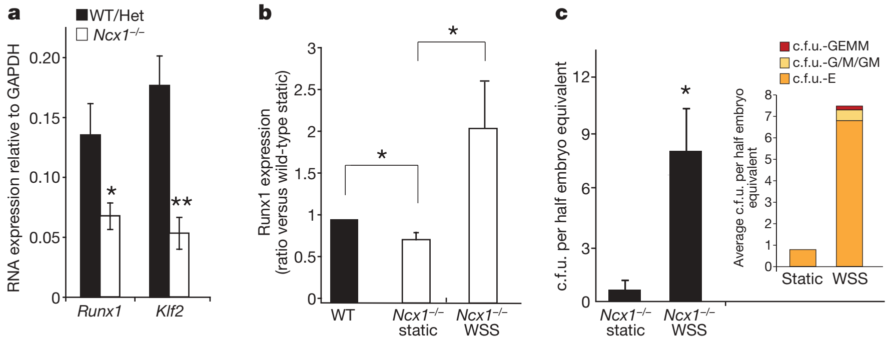


Figure 4 | Runx1 expression and c.f.u. activity are shear-stress-dependent. a, PSp isolated from E9.25 *Ncx1*$^{-/-}$ embryos show reduced gene expression levels of the haematopoietic markers *Runx1* ($P = 0.02$) and *Klf2* ($P = 0.003$) when compared to matched wild-type (WT) or heterozygous (Het) littermate controls; $n = 13$ *Ncx1*$^{-/-}$, $n = 15$ controls. **b**, Shear stress increases the expression of Runx1 in E9.25 *Ncx1*$^{-/-}$ PSp cultures to levels comparable to the one observed in littermate controls; real-time quantitative PCR, $n = 4$, ANOVA, $P = 0.03$. Post-hoc multiple comparisons are two-tailed *t*-test with $P < 0.02$. **c**, Shear stress induces c.f.u. activity in *Ncx1*$^{-/-}$ PSp-derived cells. Average ± s.e.m.; $P = 0.01$. Inset shows average distribution of haematopoietic colony types at day 14 after replating; $n = 6$. $^{*}P < 0.05$, $^{**}P < 0.005$.

cells derived from the PSp region of *Ncx1*$^{-/-}$ embryos, and indicates that the haematopoietic deficiency of *Ncx1*$^{-/-}$ embryos is not a cell-autonomous defect but rather a consequence of alterations in haemodynamics affecting vascular haemogenic sites.

Our results demonstrate that biomechanical forces stimulate embryonic haematopoiesis both in ES cell cultures and within murine embryos. These data establish a link between the initiation of the heartbeat, vascular flow and embryonic haematopoietic development, and provide new perspectives for the manipulation and production of haematopoietic progenitors from pluripotent stem cells *in vitro*, a key process for the implementation of stem-cell-based therapy of haematological diseases.

METHODS SUMMARY

Haemodynamic shear stress in the embryonic aorta was estimated using previously published fluid dynamic data from the developing mouse embryo[10,17,30] and assuming circular Poiseuille flow. Ainv15 CDX4-inducible mouse embryonic stem cells (iCDX4 ES) were cultured as described[22] and differentiated via the hanging-drop embryoid body method. Differentiated ES cells were plated at 100,000 per cm^2 on 95 cm^2 plates, and exposed to shear stress in the presence or absence of 2 mM N(G)-nitro-L-arginine methyl ester (L-NAME) or 2 mM N(G)-nitro-D-arginine methyl ester (D-NAME). Pregnant Swiss–Webster mice were purchased from Taconic farms. Noon on the day the copulation plug was found was designated 0.5 days of gestation (E0.5). PSp/AGM regions were dissected with a conservative approach, preserving the somites. Two-dimensional primary PSp/AGM cultures were optimized in the absence or presence of growth factors (ECGS, heparin, SCF, VEGF, TPO and Flt3 ligand).

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 18 April 2008; accepted 23 April 2009.
Published online 13 May 2009.

1. Hove, J. R. *et al.* Intracardiac fluid forces are an essential epigenetic factor for embryonic cardiogenesis. *Nature* **421**, 172–177 (2003).
2. Lucitti, J. L. *et al.* Vascular remodeling of the mouse yolk sac requires hemodynamic force. *Development* **134**, 3317–3326 (2007).
3. Garcia-Porrero, J. A., Godin, I. E. & Dieterlen-Lievre, F. Potential intraembryonic hemogenic sites at pre-liver stages in the mouse. *Anat. Embryol.* **192**, 425–435 (1995).
4. North, T. E. *et al.* Runx1 expression marks long-term repopulating hematopoietic stem cells in the midgestation mouse embryo. *Immunity* **16**, 661–672 (2002).
5. Rhodes, K. E. *et al.* The emergence of hematopoietic stem cells is initiated in the placental vasculature in the absence of circulation. *Cell Stem Cell* **2**, 252–263 (2008).
6. Lensch, M. W. & Daley, G. Q. Origins of mammalian hematopoiesis: *in vivo* paradigms and *in vitro* models. *Curr. Top. Dev. Biol.* **60**, 127–196 (2004).
7. Mikkola, H. K., Fujiwara, Y., Schlaeger, T. M., Traver, D. & Orkin, S. H. Expression of CD41 marks the initiation of definitive hematopoiesis in the mouse embryo. *Blood* **101**, 508–516 (2003).
8. Garcia-Cardena, G. *et al.* Dynamic activation of endothelial nitric oxide synthase by Hsp90. *Nature* **392**, 821–824 (1998).
9. Haar, J. L. & Ackerman, G. A. A phase and electron microscopic study of vasculogenesis and erythropoiesis in the yolk sac of the mouse. *Anat. Rec.* **170**, 199–223 (1971).
10. Ji, R. P. *et al.* Onset of cardiac function during early mouse embryogenesis coincides with entry of primitive erythroblasts into the embryo proper. *Circ. Res.* **92**, 133–135 (2003).
11. Tavian, M. *et al.* The vascular wall as a source of stem cells. *Ann. NY Acad. Sci.* **1044**, 41–50 (2005).
12. Garin, G. & Berk, B. C. Flow-mediated signaling modulates endothelial cell phenotype. *Endothelium* **13**, 375–384 (2006).
13. Kabrun, N. *et al.* Flk-1 expression defines a population of early embryonic hematopoietic precursors. *Development* **124**, 2039–2048 (1997).
14. Lengerke, C. *et al.* BMP and WNT specify hematopoietic fate by activation of the CDX-Hox pathway. *Cell Stem Cell* **2**, 72–82 (2008).
15. Medvinsky, A. & Dzierzak, E. Definitive hematopoiesis is autonomously initiated by the AGM region. *Cell* **86**, 897–906 (1996).
16. Cumano, A., Dieterlen-Lievre, F. & Godin, I. Lymphoid potential, probed before circulation in mouse, is restricted to caudal intraembryonic splanchnopleura. *Cell* **86**, 907–916 (1996).
17. Phoon, C. K., Aristizabal, O. & Turnbull, D. H. Spatial velocity profile in mouse embryonic aorta and Doppler-derived volumetric flow: a preliminary model. *Am. J. Physiol. Heart Circ. Physiol.* **283**, H908–H916 (2002).
18. Ku, D. Blood flow in arteries. *Annu. Rev. Fluid Mech.* **29**, 399–434 (1997).
19. Yamamoto, K. *et al.* Fluid shear stress induces differentiation of Flk-1-positive embryonic stem cells into vascular endothelial cells *in vitro*. *Am. J. Physiol. Heart Circ. Physiol.* **288**, H1915–H1924 (2005).
20. Basu, P. *et al.* KLF2 is essential for primitive erythropoiesis and regulates the human and murine embryonic beta-like globin genes *in vivo*. *Blood* **106**, 2566–2571 (2005).
21. Parmar, K. M. *et al.* Integration of flow-dependent endothelial phenotypes by Kruppel-like factor 2. *J. Clin. Invest.* **116**, 49–58 (2006).
22. Wang, Y., Yates, F., Naveiras, O., Ernst, P. & Daley, G. Q. Embryonic stem cell-derived hematopoietic stem cells. *Proc. Natl Acad. Sci. USA* **102**, 19081–19086 (2005).
23. Aicher, A. *et al.* Essential role of endothelial nitric oxide synthase for mobilization of stem and progenitor cells. *Nature Med.* **9**, 1370–1376 (2003).
24. Rees, D. D., Palmer, R. M. & Moncada, S. Role of endothelium-derived nitric oxide in the regulation of blood pressure. *Proc. Natl Acad. Sci. USA* **86**, 3375–3378 (1989).
25. Ferkowicz, M. J. *et al.* CD41 expression defines the onset of primitive and definitive hematopoiesis in the murine embryo. *Development* **130**, 4393–4403 (2003).
26. Koushik, S. V. *et al.* Targeted inactivation of the sodium-calcium exchanger (Ncx1) results in the lack of a heartbeat and abnormal myofibrillar organization. *FASEB J.* **15**, 1209–1211 (2001).
27. Lux, C. T. *et al.* All primitive and definitive hematopoietic progenitor cells emerging prior to E10 in the mouse embryo are products of the yolk sac. *Blood* **111**, 3435–3438 (2007).
28. Simeone, A., Daga, A. & Calabi, F. Expression of runt in the mouse embryo. *Dev. Dyn.* **203**, 61–70 (1995).
29. North, T. *et al.* Cbfa2 is required for the formation of intra-aortic hematopoietic clusters. *Development* **126**, 2563–2575 (1999).
30. Jones, E. A., Baron, M. H., Fraser, S. E. & Dickinson, M. E. Measuring hemodynamic changes during mammalian development. *Am. J. Physiol. Heart Circ. Physiol.* **287**, H1561–H1569 (2004).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank G. Losyev for assistance with flow cytometry, S. Schmitt for critical help in optimizing AGM culture conditions, C. Lengerke and Y. Mukouyama for critical discussions. L.A. was partially funded by the Giovanni Armenise-Harvard Foundation. O.N. was partially funded by the Barrie de la Maza Foundation. G.G.-C. was supported by grants from the National Institutes of Health and G.Q.D was supported by grants from the National Institutes of Health (NIH), and the NIH Director's Pioneer Award of the NIH Roadmap for Medical Research. G.Q.D. is a recipient of the Burroughs Wellcome Fund Clinical Scientist Award in Translational Research and is an Investigator of the Howard Hughes Medical Institute.

Author Contributions L.A., O.N, G.G.-C. and G.Q.D. conceived ideas, designed experiments, analysed results and wrote the manuscript. P.J.M. performed the haemodynamic shear stress estimation and programmed the biomechanical stimuli. L.A., O.N., P.L.W., J.G.-S., S.M.-F. and A.S.-D. performed experiments. M.W.L. conceived ideas and contributed to experimental design. M.Y. and M.C.Y. set up timed pregnancies and isolated *Ncx1* null and wild-type mouse embryos. All authors edited and reviewed the final manuscript. G.Q.D. and G.G.-C. co-directed the project.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to G.Q.D. (george.daley@childrens.harvard.edu) or G.G.-C. (guillermo_garcia-cardena@hms.harvard.edu).

METHODS

Mouse ES cell culture and differentiation. Ainv15 CDX4-inducible mouse embryonic stem cells (iCDX4 ES) were cultured on mitomycin-treated mouse embryonic fibroblasts (MEFs) (Chemicon, strain CF-1) and passaged every 2–3 days with 0.25% trypsin/EDTA as previously described[22]. To induce differentiation, ES cells were trypsinized, gently dispersed by repetitive pipetting and replated for 45 min at 37 °C for MEF depletion. Floating and loosely adherent cells were gathered, checked for viability with trypan blue, and diluted in embryoid body differentiation media (IMDM, 10% FCS Stem Cell technologies catalogue number 06952, 10 units ml^{-1} penicillin, 10 μg ml^{-1} streptomycin, 1 mM L-glutamine, 3.8 nl ml^{-1} monothioglycerol, 0.2 mg ml^{-1} Fe-saturated transferrin, 0.5 mg ml^{-1} ascorbic acid) to a final concentration of 6.7×10^3 cell ml^{-1}. Embryoid bodies were cultured via the hanging-drop method in 15-μl drops on 15 cm bacterial dishes. Typically 100 15-cm plates of embryoid bodies were used per experiment. After 48 h, embryoid bodies were pooled into calcium- and magnesium-containing warm PBS and collected by sedimentation in 50-ml Falcon tubes. Pooled embryoid bodies were transferred to 10-cm tissue culture dishes containing 10 ml of embryoid body differentiation media each, placed on an orbital shaker moving at 75 r.p.m., and incubated for an additional 30 h. After this second culture phase, day 3.25 embryoid bodies were collected by precipitation, washed two times in 30-ml calcium-free PBS and trypsinized for 1 min in a 1:1 dilution of calcium-free PBS and 0.25% Trypsin-EDTA. Trypsinization was blocked by addition of 10 ml embryoid body differentiation media.

Haemodynamic shear-stress estimation. Wall shear stress acting along the dorsal aorta transiently increases during embryonic development once the heart begins beating. This increase in shear stress may be attributed to both an increase in blood flow velocity, a result of increasing cardiac output[31,32], and an increase in viscosity, an effect of increasing haematocrit[33]. However, during the first two embryonic days after the initiation of a beating heart (E8–E10), the haematocrit rises, but remains below 20%. This change has an insignificant contribution to blood viscosity[34] and, therefore, at these early time points blood viscosity can be assumed constant and equal to 0.015 dyn s^{-1} cm^{-2}. Using ultrasound biomicroscopy (UBM) data, one study[32] has previously determined blood flow in the developing embryo dorsal aorta to be laminar throughout the cardiac cycle with minimal skewing of the spatial velocity profile, as determined by calculation of non-dimensional fluid dynamic parameters at E11.5, including the Reynolds number (157), Womersley parameter (0.63) and Dean number (39), respectively. Assuming such fluid dynamic parameters remain valid for earlier time points during development, as well as that the instantaneous blood velocity profile is parabolic, we estimated the dorsal aorta wall shear stress based on circular Poiseuille flow in which the wall shear stress, τ, is calculated by $\tau = 8\mu V/D$, where μ is the apparent viscosity (dyn s^{-1} cm^{-2}), V (mm s^{-1}) is the time-average mean blood flow velocity, and D (mm) is the vessel diameter. A similar strategy[33] has been used to estimate embryonic shear stress levels in the developing embryo (E8.5–10.5).

Mouse ES cell exposure to shear stress. Differentiated ES cells were plated at 100,000 per cm^2 on horizontal 95-cm^2 plates, compatible with a Dynamic Flow System[35]. The plates were coated with a thick layer of 1% gelatin (DIFCO 214340, Becton Dickinson). After 12 h, cells were washed three times with 10 ml of warm PBS with calcium and incubated for 45 min with embryoid body differentiation media with or without 2 mM N(G)-nitro-L-arginine methyl ester (L-NAME) or 2 mM N(G)-nitro-D-arginine methyl ester (D-NAME) (Sigma-Aldrich). Following 45 min of incubation time, cells were either moved to the Dynamic Flow System and exposed to shear stress or maintained in a standard cell culture incubator. The shear stress ramped in a step-wise fashion from 0 to 5 dyn cm^{-2} over a period of 10 h. Following the 10-h ramping period, the cells were then exposed to a constant 5 dyn cm^{-2} shear stress for 38 h. In cells exposed to shear stress, culture media with or without D-NAME or L-NAME was automatically exchanged and stored at 37 °C and 5% CO_2. In static controls, culture media was replaced after 24 h and stored in 75-cm^2 flasks in a cell culture incubator. After 48 h exposure to shear stress or growth under static conditions, cells were detached by 10 min incubation with a 1:1 dilution of Trypsin–Versene mixture (Cambrex, number 17-161E) and Hank's Balanced Salt Solution.

Primary AGM/PSp culture. E9.25/9.5 or E10.5 murine embryos were collected in PBS, and a conservative dissection of the AGM/PSp region preserving the somites was performed. Typically, 32 embryos were used per condition on either the full Dynamic Flow System cell growth surface area (95 cm^2, E10.5 embryos) or about one-third of the Dynamic Flow System cell growth surface area (38 cm^2, E9.5). For the *Ncx1*$^{-/-}$ rescue experiment, single E9.25 PSp were disaggregated independently, divided in two and plated at a density of half embryo equivalent per 0.3 cm^2 either in the Dynamic Flow System or in a 96-well format system. E9.5/E9.25 embryos were dissected within a maximum of 2.5 h and exposed to dispase (see below for details) without any delay. PSp-derived cells were plated in M5300 Myelocult medium (Stem Cell technologies) in the absence of hydrocortisone on surfaces coated for 1 h with a 1:20 matrigel dilution (Beckton Dickinson), cells adhered for 6.5 h and were then moved to the shear stress culture system for 36 h in either static condition or exposed to shear stress. When culturing cells in 96-well format, Myelocult was enriched with 25 ml l^{-1} of 1 M HEPES (Invitrogen). The shear stress ramped in a step-wise fashion from 0 to 5 dyn cm^{-2} over a period of 10 h. After the 10-h ramping period, as described in Supplementary Fig. 1a, the cells were exposed to a constant 5 dyn cm^{-2} shear stress for 26 h. Cells were harvested by dissociation with 0.1% dispase for 20 min. E10.5 AGM regions were dissociated in dispase in the same manner. E10.5 AGM-derived cells were plated overnight on 1% gelatin-coated surfaces in IMDM 10% FCS, 1× NEAA (Gibco), 1× β-mercaptoethanol (Gibco), 1 mM L-glutamine, 10 units ml^{-1} penicillin, 10 μg ml^{-1} streptomycin, 50 μg ml^{-1} ECGS, 100 μg ml^{-1} heparin, 100 ng ml^{-1} SCF, 40 ng ml^{-1} VEGF, 40 ng ml^{-1} TPO and 100 ng ml^{-1} Flt3 ligand overnight. IMDM media supplementation was then reduced to 25 μg ml^{-1} ECGS, 50 μg ml^{-1} heparin, 25 ng ml^{-1} SCF, 20 ng ml^{-1} VEGF, 10 ng ml^{-1} TPO and 25 ng ml^{-1} Flt3 ligand when cells were moved to the Dynamic Flow System and kept for 48 h in either static condition or exposed to shear stress. The shear stress ramped in a step-wise fashion from 0 to 5 dyn cm^{-2} over a period of 10 h. After the 10-h ramping period, the cells were then exposed to a constant 5 dyn cm^{-2} shear stress for 38 h. Cells were harvested as described herein for mouse ES-derived cells. Non-adherent cells were gathered by centrifugation. Embryos derived from *Ncx1*$^{+/-}$ mothers were identified as mutant (*Ncx1*$^{-/-}$) or heterozygous/wild type by visual inspection of the yolk sac vasculature. Genotype was further confirmed by automated genotyping (Transnetyx) on embryonic biopsies collected at the time of dissection. Dispase stock solution was prepared diluting the commercially available powder form (GIBCO) into PBS without calcium or magnesium to a final concentration of 2.5 g per 100 ml. The solution was filtered in 0.22 μM tissue culture filters and stored at −20 °C. The stock was diluted 1:20 in PBS with calcium and magnesium to make working solutions. Embryonic tissue was disaggregated by incubating it in dispase working solution for 40 min in a thermal mixer rotating a 1,100 r.p.m. at 37 °C. After the first 20 min of incubation, dispase solution was pipetted up and down about 15 times to break larger tissue pieces. To detach cells plated on matrigel, cells were covered in dispase working solution, incubated 20 min in a 37 °C incubator and then collected by pipetting. Before removing dispase, cells were incubated for another 20 min in a 37 °C water bath.

RNA extraction and quantitative real-time PCR. Total RNA was extracted with the QIAGEN RNAeasy Micro Kit, according to the manufacturer's instructions. For RNA extraction from CD41$^+$c-Kit$^+$ cells, approximately 7,000 cells were sorted directly into 300 μl of lysis buffer. Reverse transcription of RNA was performed using Applied Biosystems Multiscribe DNA polymerase, according to the manufacturer's instructions. When processing FACS-sorted cells, RNA was eluted in 20 μl and all collected RNA was used in 50 μl reverse transcription reactions. Real-time Taqman PCR (Applied Biosystems) was performed in 20 μl reactions with primers provided by Applied Biosystems, according to the manufacturer's instructions.

FACS sorting and analysis. Cells were incubated with Fc Block antibody (BD 553141, dilution 1:100) and then stained with anti-CD41 antibody (BD 553848, 1:150 dilution), anti-c-Kit antibody (BD 553356, 1:150 dilution), anti-Flk1 antibody (BD 555308, 1:150 dilution), anti-CD31 antibody (BD 553372, 1:150 dilution) and 7AAD (BD 559925, 1:75 dilution) and sorted with a FACS-Aria system (BD Biosciences).

Methylcellulose c.f.u. assay. Methylcellulose assay was performed as previously described[36] in M3434 complete methylcellulose (Stem Cell Technologies) and scored on day 8–10 after plating. When assessing c.f.u. activity on cells exposed to shear stress in 96-well format (*Ncx1*$^{-/-}$ PSp-derived cells), at the end of the experiment shear stress media was exchanged with 100 μl of M3434 complete methylcellulose medium. Cells were allowed to grow in M3434 in the 96-well plate for 4 days and then replated in 2 ml of complete M3434 media in optical quality 35-mm dishes. To replate, 100 ml of IMDM 15% FCS was added to the wells and cells were collected by repeated pipetting before washing the well with additional 100 μl of IMDM 15% FCS. Plates were scored on day 14 after replating.

Blockade of NO production *in vivo*. Day 7 post coitum timed-pregnant Swiss–Webster mice were purchased from Taconic Farms. Noon on the day the copulation plug was found was designated 0.5 days. On day E8.5 (day 9 post coitum) mice were intraperitoneally injected with 300 mg kg^{-1} of L-NAME or D-NAME (Sigma) dissolved in sterile, apyrogenic saline solution using a dose volume of 10 ml kg^{-1} (ref. 37). The injection was repeated after 24 h. Exactly 24 h after the second injection, mice were killed with CO_2, embryos were harvested, and a conservative dissection of the AGM was performed. Embryo trunks were dissociated with 0.1% dispase and analysed with the methylcellulose assay. Experiments were carried out with Institutional Animal Care and Use

Committee approval from Harvard Medical School or Children's Hospital, Boston.

Statistical analysis. All the statistical analyses were performed using un-paired two-tailed Student's *t*-test assuming experimental samples of equal variance, unless otherwise specified.

31. Ji, R. P. *et al.* Onset of cardiac function during early mouse embryogenesis coincides with entry of primitive erythroblasts into the embryo proper. *Circ. Res.* 92, 133–135 (2003).
32. Phoon, C. K., Aristizabal, O. & Turnbull, D. H. Spatial velocity profile in mouse embryonic aorta and Doppler-derived volumetric flow: a preliminary model. *Am. J. Physiol. Heart Circ. Physiol.* 283, H908–H916 (2002).
33. Jones, E. A., Baron, M. H., Fraser, S. E. & Dickinson, M. E. Measuring hemodynamic changes during mammalian development. *Am. J. Physiol. Heart Circ. Physiol.* 287, H1561–H1569 (2004).
34. Nosek, T. M. *Essentials of Human physiology — Cardiac and Circulatory Physiology* (Gold Standard Multimedia, 2000).
35. Blackman, B. R., Garcia-Cardena, G. & Gimbrone, M. A. Jr. A new *in vitro* model to evaluate differential responses of endothelial cells to simulated arterial shear stress waveforms. *J. Biomech. Eng.* 124, 397–407 (2002).
36. Wang, Y., Yates, F., Naveiras, O., Ernst, P. & Daley, G. Q. Embryonic stem cell-derived hematopoietic stem cells. *Proc. Natl Acad. Sci. USA* 102, 19081–19086 (2005).
37. Tiboni, G. M., Marotta, F. & Barbacane, L. Production of axial skeletal malformations with the nitric oxide synthesis inhibitor NG-nitro-L-arginine methyl ester (L-NAME) in the mouse. *Birth Defects Res. B Dev. Reprod. Toxicol.* 80, 28–33 (2007).

LETTERS

Subcellular homeostasis of phytohormone auxin is mediated by the ER-localized PIN5 transporter

Jozef Mravec[1,2], Petr Skůpa[3], Aurélien Bailly[4], Klára Hoyerová[3], Pavel Křeček[3], Agnieszka Bielach[1], Jan Petrášek[3,5], Jing Zhang[1,2], Vassilena Gaykova[2], York-Dieter Stierhof[2], Petre I. Dobrev[3], Kateřina Schwarzerová[5], Jakub Rolčík[6], Daniela Seifertová[3], Christian Luschnig[7], Eva Benková[1], Eva Zažímalová[3], Markus Geisler[4] & Jiří Friml[1,8]

The plant signalling molecule auxin provides positional information in a variety of developmental processes by means of its differential distribution (gradients) within plant tissues[1]. Thus, cellular auxin levels often determine the developmental output of auxin signalling. Conceptually, transmembrane transport and metabolic processes regulate the steady-state levels of auxin in any given cell[2,3]. In particular, PIN auxin-efflux-carrier-mediated, directional transport between cells is crucial for generating auxin gradients[2,4,5]. Here we show that *Arabidopsis thaliana PIN5*, an atypical member of the *PIN* gene family, encodes a functional auxin transporter that is required for auxin-mediated development. PIN5 does not have a direct role in cell-to-cell transport but regulates intracellular auxin homeostasis and metabolism. PIN5 localizes, unlike other characterized plasma membrane PIN proteins, to endoplasmic reticulum (ER), presumably mediating auxin flow from the cytosol to the lumen of the ER. The ER localization of other PIN5-like transporters (including the moss PIN) indicates that the diversification of PIN protein functions in mediating auxin homeostasis at the ER, and cell-to-cell auxin transport at the plasma membrane, represent an ancient event during the evolution of land plants.

PIN proteins are plant-specific integral membrane proteins directly involved in auxin efflux from cells[6], the asymmetric localization of which at the plasma membrane determines the directionality of auxin flow[7] (Supplementary Fig. 1). The *Arabidopsis* genome encodes eight PIN-related sequences, five of which have already been characterized in terms of their function in cellular auxin efflux and role in many aspects of auxin-related development[2], including embryogenesis[8], organogenesis[9,10], root meristem patterning[11,12], vascular tissue differentiation and regeneration[13–15], and tropisms[16,17]. PIN5 (At5g16530) belongs to the functionally uncharacterized subclade of PIN proteins formed by *Arabidopsis* PIN5, PIN6 and PIN8 (Fig. 1a) and that also includes PIN5-like proteins from other plant species (Supplementary Fig. 2). The typical feature of these proteins is reduction of the central hydrophilic loop that has been found in all PIN proteins characterized so far (Fig. 1b).

To test whether PIN5 functions, similarly to the conventional PIN proteins, in auxin efflux, we performed an auxin transport assay in a heterologous yeast system. PIN5 tagged with a haemagglutinin epitope (PIN5–HA) was targeted predominantly to the plasma membrane, as was confirmed by sucrose gradient fractionation (Supplementary Fig. 3a). Similarly to what was shown for representative members of the PIN family[6], PIN5 expression in yeast led to increased auxin efflux, as indicated by decreased retention of the radioactively labelled auxins indole-3-acetic acid (IAA) and α-naphthaleneacetic acid (NAA; Fig. 1c). In contrast, transport experiments using yeast microsomes demonstrated that PIN5 did not mediate the transport of the auxin conjugate IAA–aspartate (IAA–Asp; Supplementary Fig. 3b). These results show that, despite the differences in the overall protein structure and the reduced hydrophilic loop, PIN5 can mediate, as other PIN proteins[6], specific auxin export through the plasma membrane.

To analyse the PIN5 expression *in planta*, we fused a *PIN5* genomic fragment with the β-glucuronidase (*GUS*) reporter gene. GUS staining revealed increased PIN5 expression in elongating parts of hypocotyl, cotyledon vasculature and guard cells as well as weaker expression in root pericycle and root tip and at later developmental stages in leaves, stems and flowers (Supplementary Fig. 4a–f). The

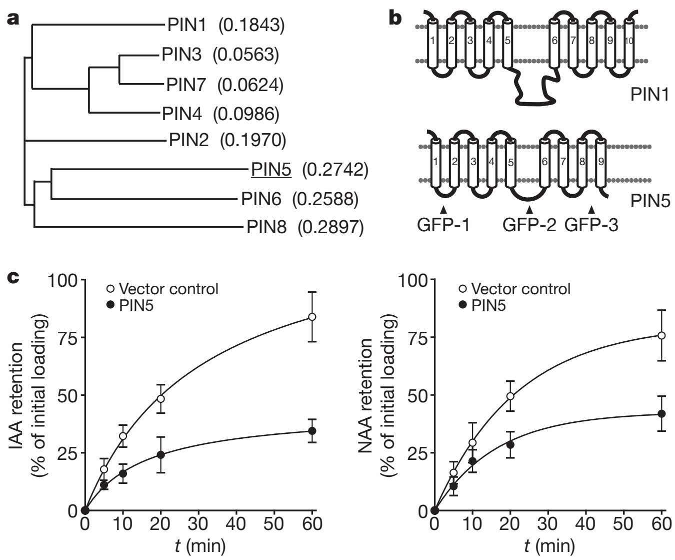


Figure 1 | PIN5 auxin transporter. a, Phylogenetic tree of the *Arabidopsis* PIN protein family with indicated distances. PIN5, PIN8 and PIN6 represent the distant subclade. **b**, Predicted organization of transmembrane domains in PIN5 reveals the reduction of the middle hydrophilic loop that is the hallmark of PIN1-type proteins. Positions of GFP insertions are indicated. **c**, In the yeast auxin efflux assay, decreased retention of IAA and NAA in the PIN5-expressing yeast implies an auxin-export function for PIN5. Error bars represent s.e.m.; $n=5$.

[1]Department of Plant Systems Biology, VIB and Department of Plant Biotechnology and Genetics, Ghent University, 9052 Gent, Belgium. [2]Center for Plant Molecular Biology (ZMBP), University Tübingen, D-72076 Tübingen, Germany. [3]Institute of Experimental Botany, ASCR, 165 02 Praha 6, Czech Republic. [4]Zürich-Basel Plant Science Center, University of Zurich, Institute of Plant Biology, Molecular Plant Physiology, CH-8008 Zurich, Switzerland. [5]Department of Plant Physiology, Faculty of Science, Charles University, 128 44 Praha 2, Czech Republic. [6]Laboratory of Growth Regulators of the Institute of Experimental Botany, ASCR and Faculty of Science, Palacký University, 783 71 Olomouc, Czech Republic. [7]Department for Applied Genetics and Cell Biology, University of Natural Resources and Applied Life Sciences - BOKU, A-1190 Wien, Austria. [8]Department of Functional Genomics and Proteomics, Faculty of Science, Masaryk University, 625 00 Brno, Czech Republic.

PIN5 expression pattern was confirmed by *in situ* hybridization and quantitative reverse-transcription polymerase chain reaction (qRT–PCR; Supplementary Fig. 4g, h) and is consistent with publicly available microarray data (https://www.genevestigator.ethz.ch/at/). In contrast to other tested PINs, which are typically upregulated in response to auxin[18], PIN5 was transcriptionaly downregulated after IAA treatment (Supplementary Fig. 4i). Together, these findings indicate that PIN5 acts as an auxin-regulated broadly expressed auxin transport protein.

To analyse the role of PIN5 in plant development, we characterized three T-DNA insertion mutants in which the *PIN5* locus was fully disrupted (Fig. 2a, b). All *pin5* mutant alleles showed pronounced defects in lateral root initiation as well as in root and hypocotyl growth (Fig. 2c, d). Furthermore, *pin5* root growth was less sensitive to low concentrations of IAA, whereas its sensitivity to auxin analogues and auxin transport inhibitors was unchanged (Fig. 2e and Supplementary Fig. 5a). To get further insights into the role of PIN5, we generated gain-of-function alleles. We used a two-component UAS/GAL4-based system, driving the PIN5-*myc* overexpression by a strong embryonic and meristematic promoter in *RPS5A* ≫ *PIN5-myc* (*PIN5 OX*). *PIN5 OX* seedlings were defective in root and hypocotyl growth; cotyledons did not fully expand and were epinastic (Fig. 2f and Supplementary Fig. 5b). The less affected *PIN5 OX* adult plants formed narrow, wrinkled leaves and bushy inflorescences (Fig. 2g, h), whereas the more affected primary transformants had small filamentous leaves and failed to flower (Fig. 2g, inset). Initiation of lateral

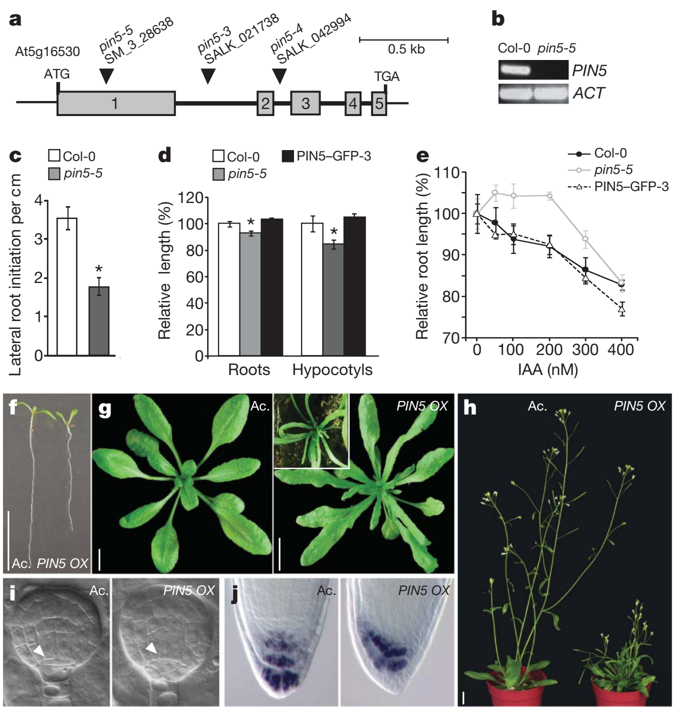


Figure 2 | Phenotypes of *pin5* knockout and gain-of-function mutants. **a**, Structure of the *PIN5* gene with five exons (numbered) and depiction of three T-DNA insertions characterizing the *pin5* loss-of-function mutations. **b**, *PIN5* mRNA is not detected in the *pin5-5* mutant by RT–PCR. Actin (ACT) gene was used as a reference control. **c–e**, Developmental defects in *pin5* mutants. *pin5-5* mutants initiate fewer lateral roots (**c**), show defects in hypocotyl and root elongation (**d**), and root growth is resistant to lower concentrations of IAA (**e**). Error bars represent s.e.m., $n = 20$, $*P < 0.05$. **f–j**, Developmental defects of the *PIN5 OX* line in comparison to the *RPS5A:GAL4* activator line (Ac.). *PIN5 OX* seedlings show epinastic, not fully opened cotyledons, shorter roots (**f**) and defects in organization of root meristem as visualized by lugol staining (**j**); adult plants have wrinkled and narrow rosette leaves (inset shows a strongly affected line with filamentous leaves; **g**) and dwarfed and bushy inflorescences (**h**); embryos show aberrations in hypophysis division (indicated by arrowheads; **i**). Scale bars, 1 cm.

roots occurred normally, but the development of primordia progressed more slowly during the later stages than in the control activator line (Supplementary Fig. 5c, d). *PIN5 OX* embryos that overexpressed PIN5 from early embryogenesis onwards were also defective in the establishment of the root pole ($n = 8$ out of 62; Fig. 2i) and postembryonically in differentiation and organization of columella ($n = 7$ out of 25; Fig. 2j). In summary, analysis of the *pin5* loss- and gain-of-function mutant phenotypes indicates a role for PIN5 in a broad range of developmental processes, mainly in those that are regulated by auxin.

To study the involvement of PIN5 in auxin-mediated development further, we examined the activity of the synthetic auxin-responsive reporter *DR5rev::GFP* in *pin5* mutants. The spatial pattern of the DR5 activity did not change markedly but the overall DR5 activity in the root meristems was increased in *pin5* and decreased in *PIN OX* (Supplementary Fig. 6). These results do not support a role for PIN5 in intercellular auxin transport and generation of auxin gradients as shown for the other PIN proteins[2], but instead indicate defects in auxin signalling or homeostasis. To address this issue, we directly measured auxin levels in the *pin5* mutants. Consistently, extraction and gas chromatography/mass spectrometry detection of IAA in root tips and rosette leaves showed that loss-of-function *pin5* mutants had higher levels of endogenous free auxin than the controls whereas the root tips or rosette leaves of *PIN OX* plants had decreased free IAA levels (Fig. 3a). Furthermore, the *pin5* loss-of-function allele showed an unusual genetic interaction with the *yucca* gain-of-function mutant, which overexpresses a rate-limiting enzyme in auxin biosynthesis and has increased auxin levels[19]. The additive effects of these mutations in root and the enhanced effects in hypocotyls (Fig. 3b) are not easy to interpret but further hint at an auxin-homeostasis-related role of PIN5.

To investigate how the loss and gain of the PIN5 function might influence cellular auxin levels, we measured the auxin efflux from mesophyll protoplasts isolated from rosette leaves. *pin5* protoplasts showed an increased, whereas *PIN5 OX* protoplasts had a decreased, efflux of IAA (Fig. 3c). Given the demonstrated auxin export function of PIN5 in yeast (see Fig. 1c), these observations were unexpected because, as shown for other PINs[6], loss of function of a plasma-membrane-localized auxin exporter would be expected to result in reduced export and vice versa. Thus, these data indicate that PIN5 mediates auxin homeostasis by a mechanism that is distinct from the known auxin efflux carriers.

Next we addressed the role of PIN5 in IAA metabolism. We generated cultured tobacco BY-2 cells stably transformed with a dexamethasone (DEX)-inducible PIN5 (as has been done for other PIN proteins). The induction of PIN5 expression led to rapid and notable change in the IAA metabolic profile as demonstrated by the absence of prominent peaks and the appearance of additional ones in the high-performance liquid chromatography (HPLC) profile (Fig. 3d and Supplementary Fig. 7). On the basis of the retention times, we identified the most prominent peaks as free IAA and IAA–glucosyl ester (IAA–Glc), which almost completely disappeared after PIN5 induction (Fig. 3d). Direct analysis by liquid chromatography/mass spectrometry of IAA and selected IAA conjugates confirmed the overall significant changes in rate of IAA metabolism after PIN5 induction, including a pronounced decrease in free IAA levels and an increased capacity to produce amino acid conjugates such as IAA–Asp or IAA–Glu (Fig. 3e). This unexpected finding that PIN5 regulates auxin homeostasis by modulating IAA metabolism further highlights a different mode of action for auxin transporters of the PIN5 subclade than those described for the PIN proteins characterized so far.

We then examined the subcellular localization of PIN5. We fused the *PIN5* genomic fragment with green fluorescent protein (GFP) and constitutively expressed it under the control of the 35S promoter. To minimize possible interference with targeting and protein folding, we placed GFP at three different positions within the PIN5-coding region (Fig. 1b). The expression of these constructs as exemplified by

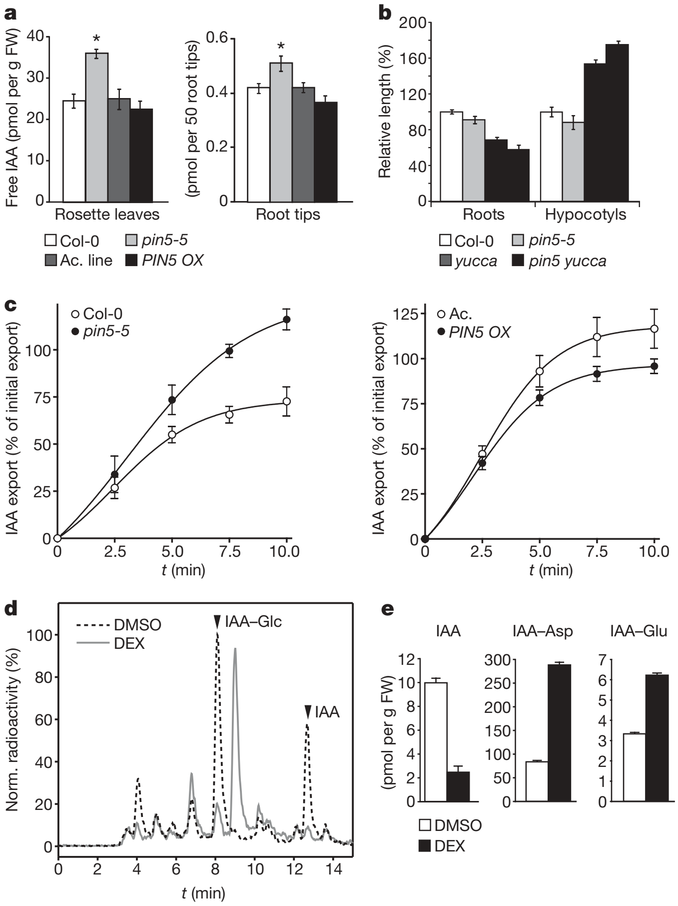


Figure 3 | PIN5 in regulation of auxin homeostasis. **a**, Changes in free IAA levels in *pin5-5* and *PIN5 OX* in rosette leaves and root tips (FW, fresh weight). Error bars represent s.e.m., $n = 3$, $*P < 0.05$. **b**, Genetic interaction between *pin5* and *yucca*. *pin5* enhances the *yucca* phenotype as shown for root and hypocotyl elongation. **c**, Protoplasts of *pin5-5* mutant exhibit a higher, whereas *PIN5 OX* a lower, rate of IAA export. Error bars represent s.e.m., $n = 5$. **d**, The HPLC chromatogram of IAA metabolic profile changes noticeably after induction of PIN5 expression in tobacco BY-2 cells (24 h of DEX induction, 2.5 h incubation with ^{3}H-IAA). **e**, Liquid chromatography/mass spectrometry detected lower levels of free IAA and strongly increased capacity to conjugate IAA to amino acids (IAA–Asp, IAA–Glu) after induction of PIN5 expression in BY-2 cells (24 h of DEX induction). Error bars represent s.e.m.

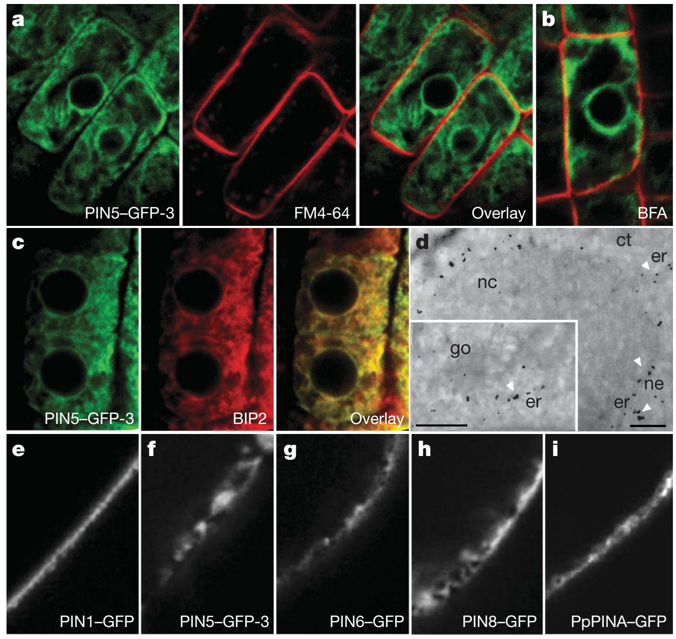


Figure 4 | PIN5 localization at the ER. **a**, PIN5–GFP-3 (green) localizes not to the plasma membrane (briefly stained by FM4-64, red) but to intracellular structures. **b**, PIN5 localization is insensitive to trafficking inhibitor brefeldin A (BFA). **c**, Complete co-localization of PIN5–GFP-3 (green) and the ER marker BIP2 (red) by immunocytochemistry. **d**, Anti-GFP-immunogold labelling on ultrathin cryosections, confirming the presence of PIN5–GFP-3 at the ER (marked by arrowheads). Abbreviations: ct, cytoplasm; er, endoplasmic reticulum; go, Golgi network; n, nucleus; ne, nuclear envelope. Scale bars, 0.5 μm. **e–i**, Localization of different PIN–GFP fusion proteins in tobacco BY-2 cells. Magnification of the plasma membrane regions is shown. PIN1–GFP at the plasma membrane (**e**). No plasma membrane but predominant ER localization of PIN5 (**f**), PIN6 (**g**), PIN8 (**h**) and PpPINA (**i**).

PIN5–GFP-3 fully complemented all aspects of *pin5* knockout phenotypes (Fig. 2d, e and Supplementary Fig. 5c). In contrast to the localization of other PIN proteins[2], all three PIN5–GFP variants as well as the PIN5-*myc* in the *PIN5 OX* plants did not show any plasma membrane but a strong intracellular signal (Fig. 4a and Supplementary Fig. 8a, b) that was, unlike that of other PIN proteins[17,20,21], insensitive to treatment with the vesicle trafficking inhibitor brefeldin A (Fig. 4b). Reliable co-localization with different ER markers such as BIP2 (spinach), Sec12 (*Arabidopsis*) and ER tracker dye (Fig. 4c and Supplementary Fig. 8c–f) as well as immunogold labelling on ultrathin cryosections (Fig. 4d) showed that PIN5 localizes to the ER. Taken together, these data indicate that the ER membrane is the site of the PIN5 action and not the plasma membrane as was shown for other members of the PIN family[2]. This also indicates that the function of PIN5 in export of auxin from the yeast cells (Fig. 1c) is due to ectopic localization of PIN5 at the plasma membrane in the heterologous system. Given the export function at the plasma membrane in yeast, the necessary inward transport orientation of the protein at ER membranes *in planta* implies a role for PIN5 in transporting auxin from the cytosol into the lumen of the ER (see model in Supplementary Fig. 1). Such PIN5-dependent transport would limit the availability of cytosolic auxin for auxin exporters at the plasma membrane and increase the IAA amount in the ER pool, where some enzymes in IAA metabolism are compartmentalized[3,22]. The defect in this mechanism in the *pin5* mutant would lead to higher levels of free IAA and increased auxin efflux, whereas PIN5 gain-of-function at the ER would have opposite effects, such as lower free IAA levels and decreased cellular auxin efflux, exactly as observed (see Fig. 3a, c, e).

Interestingly, the localization analysis in tobacco BY-2 cells showed that all *Arabidopsis* PIN5-like proteins (PIN5, PIN6 and PIN8) do not localize to the plasma membrane (as do PIN1-type proteins) but localize to the ER (Fig. 4e–h and Supplementary Fig. 9). These observations indicate that the PIN5-type proteins found throughout all land plant genomes presumably have a function at the ER. The relatively high sequence homology between ER- and plasma-membrane-localized PIN proteins indicates that only small motif changes were involved during diversification of these two classes, as shown, for example, for similar diversification of the ER- and plasma-membrane-based subclasses of plant aquaporins, which differ in a diacidic motif known to be required for exit from the ER[23]. Within the PIN family, the sequence around a putative tyrosine motif, which is a signal for recruitment into the clathrin-coated vesicles[24], consistently differs between the PIN1- and PIN5-like subclasses and might be involved in targeting diversification (Supplementary Fig. 10).

The localization data together with phylogenetic analyses revealed that the PIN protein family comprises two distinct subclasses, the plasma-membrane-localized PIN1-type and the ER-localized PIN5-type, that got separated during evolution well before the establishment of gymnosperms and angiosperms (Supplementary Fig. 2). To address which of these PIN types is ancestral, we analysed the most ancestral PIN-like sequences in the genome of the moss *Physcomitrella patens*.

The *Physcomitrella* genome[25] contains three closely related PIN sequences that cannot be conspicuously allocated to either of the two PIN subclasses based on the protein structure or sequence similarities. Nonetheless, GFP-tagged *P. patens* PINA (XP_001753612), the typical member of the *P. patens* PIN family, localizes to ER in BY-2 cells (Fig. 4i and Supplementary Fig. 9g). This indicates that *Physcomitrella* PINs are functionally related to the ER-localized PIN5-like proteins of higher plants. Thus, the ER-based role of the PIN proteins appeared very early in the evolution of land plants and might represent the ancestral function of this family of auxin transporters.

Our analysis of PIN5, an ER-localized member of the PIN auxin efflux protein family, hints at the previously unknown regulation of the cellular homeostasis of the plant signalling molecule auxin by subcellular compartmentalization (Supplementary Fig. 1). The localization as well as genetic and physiological data support a view in which PIN5 transports auxin intracellularly, from the cytoplasm into the lumen of the ER, thus reducing the auxin availability for plasma-membrane-based auxin efflux and possibly also for nucleus-located TIR1-based auxin perception[26,27]. Importantly, manipulation of the PIN5-dependent cytosol-to-ER auxin transport can markedly change the metabolic fate of IAA, suggesting a subcellular compartmentalization of auxin metabolic pathways. The importance of the ER in auxin biology has already been highlighted by the presence of components of auxin metabolism and signalling such as several de-conjugation enzymes[3,22] and the auxin-binding protein ABP1 that is essential for plant growth[28,29]. Thus, PIN5-mediated auxin uptake into the ER lumen might have a relevant role in the control of the ER-resident auxin metabolism, as well as ABP1-mediated auxin perception in this compartment. Moreover, this mode of PIN action at the ER seems to be evolutionarily ancestral as signified by the ER-localized PIN proteins in moss. Auxin as a signalling molecule and thus its homeostasis regulation has been implicated already in unicellular algae, but intercellular polar auxin transport was detected only in the moss sporophyte and evolutionary younger species[30]. Thus, it is conceivable that the ER-based PIN function in regulating the auxin homeostasis is evolutionarily older and that the plasma-membrane-based PIN function in cell-to-cell transport has been acquired later with the evolution of auxin-distribution-based patterning of multicellular plants.

METHODS SUMMARY

Material and growth conditions. For all experiments, we used *Arabidopsis thaliana* of ecotype Columbia 0 (Col-0). Insertion mutant lines were *pin5-3* (SALK_021738), *pin5-4* (SALK_042994) and *pin5-5* (SM_3_28638). Insertion sites were verified, homozygous lines selected and the absence of the *PIN5* transcript shown by RT–PCR. For detailed description of mutant, transgenic lines and DNA constructs, see Methods. Seeds were sterilized with chlorine gas and stratified at 4 °C for 2 days in the dark. Seedlings were grown vertically on half Murashige and Skoog medium supplemented with 1% sucrose and the respective drugs. Plants were grown under the stable long-day (16 h light/8 h dark) conditions at 20–22 °C.

Phenotype analyses, GUS staining and auxin measurements. Plates were scanned on a flat-bed scanner and root and hypocotyl lengths were measured with the ImageJ (http://rsb.info.nih.gov/ij/) software. Lateral root initiation, developmental stage progression analyses and GUS staining were performed as described[9]. Embryos were analysed as described[8]. Details of the free IAA measurements, metabolic profiling, yeast and protoplast transport assays and analyses on BY-2 cells are described in Methods.

Microscopy. Details regarding immunological staining, antibodies and dilutions can be found in Methods. Immunogold labelling on ultrathin cryosections was performed as described[20]. GFP samples were scanned without fixation. Confocal imaging was performed on a Leica SP2 and Zeiss LSM 5 DUO confocal microscopes.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 17 February; accepted 9 April 2009.
Published online 7 June 2009.

1. Vanneste, S. & Friml, J. Auxin: trigger of change in plant development. *Cell* **136,** 1005–1016 (2009).
2. Vieten, A., Sauer, M., Brewer, P. B. & Friml, J. Molecular and cellular aspects of auxin-transport-mediated development. *Trends Plant Sci.* **12,** 160–168 (2007).
3. Woodward, A. W. & Bartel, B. Auxin: regulation, action, and interaction. *Ann. Bot. (Lond.)* **95,** 707–735 (2005).
4. Kramer, E. M. & Bennett, M. J. Auxin transport: a field in flux. *Trends Plant Sci.* **11,** 382–386 (2006).
5. Bandyopadhyay, A. *et al.* Interactions of PIN and PGP auxin transport mechanisms. *Biochem. Soc. Trans.* **35,** 137–141 (2007).
6. Petrášek, J. *et al.* PIN proteins perform a rate-limiting function in cellular auxin efflux. *Science* **312,** 914–918 (2006).
7. Wiśniewska, J. *et al.* Polar PIN localization directs auxin flow in plants. *Science* **312,** 883 (2006).
8. Friml, J. *et al.* Efflux-dependent auxin gradients establish the apical-basal axis of *Arabidopsis. Nature* **426,** 147–153 (2003).
9. Benková, E. *et al.* Local, efflux-dependent auxin gradients as a common module for plant organ formation. *Cell* **115,** 591–602 (2003).
10. Reinhardt, D. *et al.* Regulation of phyllotaxis by polar auxin transport. *Nature* **426,** 255–260 (2003).
11. Friml, J. *et al.* AtPIN4 mediates sink driven auxin gradients and root patterning in *Arabidopsis. Cell* **108,** 661–673 (2002).
12. Blilou, I. *et al.* The PIN auxin efflux facilitator network controls growth and patterning in *Arabidopsis* roots. *Nature* **433,** 39–44 (2005).
13. Scarpella, E., Marcos, D., Friml, J. & Berleth, T. Control of leaf vascular patterning by polar auxin transport. *Genes Dev.* **20,** 1015–1027 (2006).
14. Xu, J. *et al.* A molecular framework for plant regeneration. *Science* **311,** 385–388 (2006).
15. Sauer, M. *et al.* Canalization of auxin flow by Aux/IAA-ARF-dependent feedback regulation of PIN polarity. *Genes Dev.* **20,** 2902–2911 (2006).
16. Luschnig, C., Gaxiola, R. A., Grisafi, P. & Fink, G. R. EIR1, a root-specific protein involved in auxin transport, is required for gravitropism in *Arabidopsis thaliana. Genes Dev.* **12,** 2175–2187 (1998).
17. Friml, J., Wiśniewska, J., Benková, E., Mendgen, K. & Palme, K. Lateral relocation of auxin efflux regulator PIN3 mediates tropism in *Arabidopsis. Nature* **415,** 806–809 (2002).
18. Vieten, A. *et al.* Functional redundancy of PIN proteins is accompanied by auxin-dependent cross-regulation of PIN expression. *Development* **132,** 4521–4531 (2005).
19. Zhao, Y. *et al.* A role for flavin monooxygenase-like enzymes in auxin biosynthesis. *Science* **291,** 306–309 (2001).
20. Geldner, N., Friml, J., Stierhof, Y.-D., Jürgens, G. & Palme, K. Auxin transport inhibitors block PIN1 cycling and vesicle trafficking. *Nature* **413,** 425–428 (2001).
21. Kleine-Vehn, J. *et al.* ARF GEF-dependent transcytosis mechanism for polar delivery of PIN auxin carriers in *Arabidopsis. Curr. Biol.* **18,** 526–531 (2008).
22. Bartel, B. & Fink, G. R. ILR1, an amidohydrolase that releases active indole-3-acetic acid from conjugates. *Science* **268,** 1745–1748 (1995).
23. Maeshima, M. & Ishikawa, F. ER membrane aquaporins in plants. *Pflugers Arch.* **456,** 709–716 (2008).
24. Ohno, H. *et al.* Interaction of tyrosine-based sorting signals with clathrin-associated proteins. *Science* **269,** 1872–1875 (1995).
25. Rensing, S. A. *et al.* The *Physcomitrella* genome reveals evolutionary insights into the conquest of land by plants. *Science* **319,** 64–69 (2008).
26. Kepinski, S. & Leyser, O. The *Arabidopsis* F-box protein TIR1 is an auxin receptor. *Nature* **435,** 446–451 (2005).
27. Dharmasiri, N., Dharmasiri, S. & Estelle, M. The F-box protein TIR1 is an auxin receptor. *Nature* **435,** 441–445 (2005).
28. Jones, A. M. & Herman, E. M. KDEL-containing auxin-binding protein is secreted to the plasma membrane and cell wall. *Plant Physiol.* **101,** 595–606 (1993).
29. Chen, J., Ullah, H., Young, J. C., Sussman, M. R. & Jones, A. M. ABP1 is required for organized cell elongation and division in *Arabidopsis* embryogenesis. *Genes Dev.* **15,** 902–911 (2001).
30. Fujita, T. *et al.* Convergent evolution of shoots in land plants: lack of auxin polar transport in moss shoots. *Evol. Dev.* **10,** 176–186 (2008).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank D. G. Robinson, J. Clarke and NASC for providing material; B. Scheres and H. Höfte for sharing unpublished data, E. Bellinvia, M. Pařezová, B. Pešek and V. Vincenzetti for technical help and M. De Cock for critical reading and help with preparing the manuscript. This work was supported by the Volkswagenstiftung (J.F., J.M., V.G.), the Odysseus programme of the FWO (J.M., J.Z.), GA ASCR IAA601630703 (J.F.) and KJB600380604 (J.P., K.S.), the Ministry of Education, Youth and Sports of the Czech Republic, project number LC06034 (E.Z., P.S., P.K., K.H., J.P., D.S.) and MSM6198959216 (J.R.), a European Research Council starting independent research grant to E.B. (A.Bi.), the Novartis Foundation (M.G.), the Swiss National Funds (M.G.), the Forschungskredit of the University of Zurich (A.Ba.) and the Austrian Science Fund (C.L).

Author Contributions J.F. and J.M. designed the research. J.M. cloned the constructs, established the mutants and performed the expression and localization analyses, P.S. designed and performed metabolic profiling, and established transgenic tobacco lines, A.Ba. and M.G. performed the transport assays in yeast

and protoplasts, K.H. performed the free IAA measurements, P.I.D. and J.R. ran HPLC and MS auxin analyses, J.P. and K.S. analysed localization of PINs in BY-2 tobacco cells, D.S. prepared PIN6-GFP construct, A.Bi. and E.B. performed the analysis of the lateral root phenotypes, J.Z. performed the site-directed mutagenesis, V.G. performed the phenotype analyses, Y.-D.S. performed the electron microscopy, C.L. originally identified the *PIN5* gene, isolated the *PIN5* cDNA and generated overexpression lines, J.F. performed the *in situ* hybridization, P.K. constructed the cladogram, and J.M., E.Z. and J.F. wrote the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to J.F. (jiri.friml@psb.vib-ugent.be).

METHODS

DNA constructs, transformation and plant material. SALK lines were obtained from the Nottingham *Arabidopsis* Stock Center. The EXOTIC line SM_3_28638 was a gift from J. Clarke. The pPIN5:PIN5-GUS was generated by amplification of the *PIN5* genomic fragment and cloning to the BamHI and NcoI sites of pCAMBIA1391Z (http://www.cambia.org/daisy/cambia/-materials/vectors/585.html#dsy585_table) and transformed to Col-0 plants. For the *PIN5–GFP* constructs, *eGFP* was introduced into genomic *PIN5* at the following positions of the predicted coding region: GFP-1, 37 (RD/QC), GFP-2, 164 (NI/SD), GFP-3, 319 (YG/LH). For *PIN8−GFP*, *eGFP* was positioned at amino acid 203 (TR/SV). *PIN6− GFP* was generated by insertion of *GFP* into the whole *PIN6* genomic fragment (nucleotides −1,876 to 4,194 from ATG) after amino acid position 401. The resulting fusions were cloned to pGREENII-35S-tNos[31] (www.pgreen.ac.uk/).

For GFP fusion constructs with *Physcomitrella* proteins, the corresponding coding genomic fragments were amplified and inserted with the Gateway recombination, between the 35S promoter and *GFP* in the vector p7FWG2 (ref. 32). Mutated PIN1–GFP was generated by replacing NPNTY with NSLSL using the Invitrogen QuikChange II XL kit. The resulting constructs were transformed to *pin5-5* or Col-0 plants. For PIN5 overexpression, a myc-tagged *PIN5* genomic fragment was cloned into pTA7002 (ref. 33) to create pUAS:PIN5-*myc*. This plasmid was used to transform the *RPS5A:GAL4* activator line[34] to create the *RPS5A* $\gg$ *PIN5-myc* line (*PIN5 OX*). For heterologous expression in yeast, the *PIN5* cDNA was fused to the HA by primer-extension PCR and cloned to the pNEV yeast expression vector, and transformed to *Saccharomyces cerevisiae* strains JK93 (ref. 35). The standard floral dip method was used for all *Arabidopsis* transformations.

BY-2 cell lines and biolistic transformation. The *GVG-PIN5* BY-2 line was created by stable transformation of pUAS:PIN5-myc (in pTA7002) to tobacco Bright Yellow-2 (BY-2) cells[36]. The methods for stable transformation, growth conditions and analyses were as described[6]. Three to five-day-old BY-2 cells were used for biolistic transformation (Helios Gene Gun, Bio-Rad) according the manufacturer's instructions. After bombardment, transformed cells were cultivated at 27 °C in the dark and analysed after 10–16 h.

RNA extraction and real-time qRT–PCR. Whole RNA of 7-day-old seedlings and different plant organs was extracted with the Qiagen RNA extraction kit and used for cDNA synthesis (SuperscriptII, Stratagene). The whole coding region of the *PIN5* gene was amplified by standard PCR (QP5_FW1: TAATCTT AGGCTATGGCTCTG and QP5_REV1: GACCGTGACTATGATGACC). Primers specific for tubulin or actin were used as a control. For auxin induction assays, 7-day-old seedlings were transferred to fresh plates with and without IAA (5 μM) and incubated for 5 h before RNA extraction. qPCR analysis was done on an iCycler (Bio-Rad, USA) with the Platinum SYBR Green qPCR Super-UDG kit (Invitrogen) as described[18]. Transcript levels were normalized to expression of MYB88A (At2g02820).

IAA determination, protoplast and yeast auxin transport assays. To determine internal levels of endogenous auxin in *Arabidopsis* leaves, free IAA was extracted by methanol/formic acid/water (15/1/4, v/v/v) from leaf rosettes of 23-day-old *Arabidopsis* plants homogenized in liquid nitrogen. The extract was purified with the dual-mode solid-phase extraction method as described[37]. Free IAA was purified on two different HPLC columns and determined on Finnigan Polaris Q GC-MS/MS. Internal levels of endogenous auxin in *Arabidopsis* root tips were determined by collecting root tips (3–5 mm) of 8-day-old *Arabidopsis* seedlings in a 1.5% sucrose solution. For extraction of free IAA from the tips, two volumes of methanol were added to one volume of the solution. After disintegration by sonication for 5 min, root tips were separated by centrifugation. Free IAA released to the solution was determined on the Finnigan Polaris Q GC-MS/MS. For measurement of IAA and its conjugates, the homogenate from BY-2 cells was purified by solid-phase extraction followed by immunoaffinity extraction. IAA and its conjugates were then quantified by HPLC (UPLC Acquity, Waters) coupled to tandem mass detection (Quatro micro API tandem quadrupole mass spectrometer; Waters). Auxin protoplast efflux and yeast loading assays were performed as described[35].

HPLC metabolic profiling. Twenty-four hours after induction, cells were prepared for the experiment by equilibration in uptake buffer without auxin as described[6]. Experiments were done in the same buffer at room temperature and in otherwise standard culture conditions. Cells were incubated with addition of 20 nM ^{3}H-IAA (specific activity 20 Ci mmol^{-1}, American Radiolabeled Chemicals) for a period of 2.5 h. Extraction and purification of auxin metabolites were performed as described[37]. The radioactive metabolites of ^{3}H-IAA were separated on HPLC using two methods. The first one was a pilot method with lower resolution. For this method was used the column Luna C8(2), 50 × 4.6 mm, 3 μm (Phenomenex) and mobile phase A: 40 mM $HCOONH_4$, pH 3.0. The second method was to a higher resolution with column Luna C18(2), 150 × 4.6 mm, 3 μm and mobile phase A: 40 mM CH_3COONH_4, pH 4.0. Mobile phase B was CH_3CN/CH_3OH, 1/1, v/v. Flow rate was 0.6 ml min^{-1} with linear gradient 30–50% B for 10 min, 50–100% B for 1 min, 100% B for 2 min, 100–30% B for 1 min. The column eluate was monitored by a Ramona 2000 flow-through radioactivity detector (Raytest) after online mixing with three volumes (1.8 ml min^{-1}) of liquid scintillation cocktail (Flo-Scint III, Packard BioScience Co.). The radioactive metabolites of ^{3}H-IAA were identified on the basis of comparison of their retention times with authentic standards.

IAA-conjugate uptake into membrane vesicles. Yeast microsomal fractions were isolated from enzymatically prepared spheroblasts from *S. cerevisiae* strain JK93da and from strains expressing pNEV (vector control) and pNEV-PIN5-HA (PIN5–HA), and uptake of [^{3}H]N-(3-IAA)-DL-aspartate (IAA–Asp; Moravek Biochemicals (1 Ci mmol^{-1})) and [14,15,19,20-^{3}H(N)]–Leukotriene C4 (LTC4; NEN Radiochemicals (131 Ci mmol^{-1})) was measured at 30 °C by using an ATP-regenerating system and the rapid filtration technique as described[38]. Uptake rates were calculated by subtracting the radioactivity measured after 1 min of incubation from corresponding 10-min values. Conditions were repeated with three independent microsome preparations, each performed in triplicate.

Yeast membrane fractionation and western blot. Total microsomes of yeast expressing PIN5–HA were separated via continuous sucrose gradient centrifugation[39], subjected to 4–20% PAGE (Long Life Gels, Life Therapeutics), and western blots were immunoprobed with anti-HA (Roche) and the plasma membrane marker anti-H^{+}-ATPase[39].

Whole-mount *in situ* hybridization, immunolocalization and lifetime confocal microscopy. Whole-mount *in situ* hybridization of *PIN5* in seedlings was performed as described[40]. As a probe, digoxigenin-labelled antisense *PIN5* cDNA was used. Whole-mount immunological staining on 4-day-old seedlings was done in an Intavis robot[41]. Antibodies were used at the following dilutions: anti-BIP2 (Hsc70), 1:200 (Stressgen Bioreagents); rabbit anti-Sec12 (ref. 42), 1:200; mouse anti-GFP, 1:600 (Roche); rabbit anti-myc, 1:600 (Sigma-Aldrich). Anti-rabbit and anti-mouse antibodies conjugated with Cy3 or fluorescein isothiocyanate (FITC; Dianova, Germany) were used at 1:600 dilutions. For FM4-64 and ER tracker dye labelling, PIN5–GFP seedlings were incubated for 3 min in the medium with a 1:1,000 dilution of FM4-64 or ER tracker dye (Invitrogen). For BFA treatment, we incubated seedlings in liquid Murashige and Skoog medium supplemented with 50 μM brefeldin A (Molecular Probes) for 1 h, after wash out and incubation for 3 min in FM4-64 to stain the plasma membrane. The intensity correlation analysis was done with ImageJ plugin as described[43,44].

Phylogenetic analyses. Protein sequences of PIN proteins were retrieved from the NCBI database with the BLAST program; the sequences from *P. patens* and *Selaginella moellendorffii* were predicted from the preliminary releases of the genomic sequences released by DOE Joint Genome Institute (www.phytozome.net). Sequences were aligned with the program MAFFT, version 5.664 (ref. 45), and the central region that contained extensive gaps was edited. Phylogenetic relationships for each protein were assessed via Bayesian inference by the computer program MrBayes version 3.1 (ref. 46). We performed Metropolis-coupled Markov chain Monte Carlo analysis of the PIN protein evolution using the following parameters: JTT+I+Γ substitution model with six categories of gamma rate, four independent models of four chains (one cold and three heated with temperature 0.2). The analysis was run for 2,000,000 generations, sampling every 100 generations; 2,000 burn-in samples were discarded. The consensus tree was deduced from the sampled trees and rooted according to the evolutionary relationships of the organisms.

31. Hellens, R. P. *et al.* pGreen: a versatile and flexible binary Ti vector for *Agrobacterium*-mediated plant transformation. *Plant Mol. Biol.* **42,** 819–832 (2000).
32. Karimi, M., Inze, D. & Depicker, A. Gateway vectors for *Agrobacterium*-mediated plant transformation. *Trends Plant Sci.* **7,** 193–195 (2002).
33. Aoyama, T. & Chua, N.-H. A glucocorticoid-mediated transcriptional induction system in transgenic plants. *Plant J.* **11,** 605–612 (1997).
34. Weijers, D., Van Hamburg, J. P., Van Rijn, E., Hooykaas, P. J. & Offringa, R. Diphtheria toxin-mediated cell ablation reveals interregional communication during *Arabidopsis* seed development. *Plant Physiol.* **133,** 1882–1892 (2003).
35. Bouchard, R. *et al.* Immunophilin-like TWISTED DWARF1 modulates auxin efflux activities of *Arabidopsis* P-glycoproteins. *J. Biol. Chem.* **281,** 30603–30612 (2006).
36. Nagata, T., Nemoto, Y. & Hasezawa, S. Tobacco BY-2 cell line as the "HeLa" cell in the cell biology of higher plants. *Int. Rev. Cytol.* **132,** 1–30 (1992).
37. Dobrev, P. & Kamínek, M. Fast and efficient separation of cytokinins from auxin and abscisic acid and their purification using mixed-mode solid-phase extraction. *J. Chromatogr. A.* **950,** 21–29 (2002).
38. Tommasini, R. *et al.* The human multidrug resistance-associated protein functionally complements the yeast cadmium resistance factor 1. *Proc. Natl Acad. Sci. USA* **93,** 6743–6748 (1996).

39. Geisler, M. *et al.* TWISTED DWARF1, a unique plasma membrane-anchored immunophilin-like protein, interacts with Arabidopsis multidrug resistance-like transporters AtPGP1 and AtPGP19. *Mol. Biol. Cell* **14**, 4238–4249 (2003).
40. Hejátko, J. *et al. In situ* hybridisation technique for mRNA detection in whole mount *Arabidopsis* samples. *Nature Protocols* **1**, 1462–1467 (2006).
41. Sauer, M., Paciorek, T., Benková, E. & Friml, J. Immunocytochemical techniques for whole-mount *in situ* protein localization in plants. *Nature Protocols* **1**, 98–103 (2006).
42. Bar-Peled, M. & Raikhel, N. V. Characterization of AtSEC12 and AtSAR1. Proteins likely involved in endoplasmic reticulum and Golgi transport. *Plant Physiol.* **114**, 315–324 (1997).
43. Li, Q. *et al.* A syntaxin 1, Gαo, and N-type calcium channel complex at a presynaptic nerve terminal: analysis by quantitative immunocolocalization. *J. Neurosci.* **24**, 4070–4081 (2004).
44. Manders, E. E. M., Verbeek, F. J. & Aten, J. A. Measurement of co-localisation of objects in dualcolour confocal images. *J. Microsc.* **169**, 375–382 (1993).
45. Katoh, K., Misawa, K., Kuma, K. & Miyata, T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* **30**, 3059–3066 (2002).
46. Ronquist, F. & Huelsenbeck, J. P. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**, 1572–1574 (2003).

LETTERS

Yurt, Coracle, Neurexin IV and the Na^+,K^+-ATPase form a novel group of epithelial polarity proteins

Patrick Laprise[1]†, Kimberly M. Lau[2], Kathryn P. Harris[1], Nancy F. Silva-Gagliardi[2], Sarah M. Paul[3], Slobodan Beronja[1]†, Greg J. Beitel[3], C. Jane McGlade[2] & Ulrich Tepass[1]

The integrity of polarized epithelia is critical for development and human health. Many questions remain concerning the full complement and the function of the proteins that regulate cell polarity[1]. Here we report that the *Drosophila* FERM proteins Yurt (Yrt)[2] and Coracle (Cora)[3] and the membrane proteins Neurexin IV (Nrx-IV)[4] and Na^+,K^+-ATPase[5] are a new group of functionally cooperating epithelial polarity proteins. This 'Yrt/Cora group' promotes basolateral membrane stability and shows negative regulatory interactions with the apical determinant Crumbs (Crb). Genetic analyses indicate that Nrx-IV and Na^+,K^+-ATPase act together with Cora in one pathway, whereas Yrt acts in a second redundant pathway. Moreover, we show that the Yrt/Cora group is essential for epithelial polarity during organogenesis but not when epithelial polarity is first established or during terminal differentiation. This property of Yrt/Cora group proteins explains the recovery of polarity in embryos lacking the function of the Lethal giant larvae (Lgl) group of basolateral polarity proteins[6,7]. We also find that the mammalian Yrt orthologue EPB41L5 (also known as YMO1 and Limulus)[8–10] is required for lateral membrane formation, indicating a conserved function of Yrt proteins in epithelial polarity.

To clarify the mechanisms of epithelial polarization, we analysed the function of the FERM-domain protein Yrt, which was shown to act as a negative regulatory component of the Crb complex in both *Drosophila* and vertebrates[2,11]. Crb regulates epithelial apical basal polarity by promoting apical membrane and apical junctional complex formation[12,13]. Crb also controls the growth of the apical membrane at late stages of epithelial differentiation[2,14]. Yrt binds to the cytoplasmic tail of Crb and restricts Crb activity in apical membrane growth[2]. However, Yrt is predominantly a basolateral protein that is recruited into the Crb complex only at late stages of epithelial differentiation[2]. Embryos completely lacking maternal and zygotic Yrt (*yrt*$^{M/Z}$) displayed polarity defects (Fig. 1a, b) before Yrt is recruited into the Crb complex at late organogenesis (stage 14; ref. 2). This raises the question whether Yrt has a function in epithelial organization as a basolateral protein.

We analysed *yrt*$^{M/Z}$ mutants and double-mutant combinations of *yrt* and genes encoding basolateral proteins for synergistic genetic interactions that could indicate functional cooperation. *yrt*$^{M/Z}$ embryos showed clear polarity defects at post-gastrula stages of development, as indicated by the basolateral mislocalization of Crb (Fig. 1b). In contrast, zygotic *yrt* mutants demonstrated only minor polarity defects in the trunk ectoderm (Fig. 1c). This indicates that *yrt* is required for polarity, and that zygotic *yrt* mutants could provide a sensitized background to reveal genetic interactions. Within the basolateral membrane, Yrt is enriched at the septate junction together with the Na^+,K^+-ATPase and other septate junction components[15] from stage 14 onwards (Supplementary Fig. 1). We found marked genetic interactions between *yrt* and *Atpα* (which encodes the α-subunit of Na^+,K^+-ATPase) and between *yrt* and *Nrx-IV*, but not between *yrt* and six other genes that encode septate junction transmembrane proteins (see Methods). In contrast to wild type (Fig. 1a) and single mutants (Fig. 1c, d, f), apical markers were mislocalized in *yrt Atpα* and *yrt Nrx-IV* double mutants similar to *yrt*$^{M/Z}$ embryos (Fig. 1b, e, g and Supplementary Fig. 2). Enhancement of the *Nrx-IV* null phenotype by *yrt* indicates that *Nrx-IV* and *yrt* have overlapping functions and do not operate in a linear pathway. The pathway relationship between *yrt* and *Atpα* remains uncertain as embryos completely devoid of Na^+,K^+-ATPase function cannot be analysed. These findings reveal previously unrecognized functions of the Na^+,K^+-ATPase and Nrx-IV in epithelial polarity.

The developmental timing of the *yrt*$^{M/Z}$ and the *yrt Atpα* and *yrt Nrx-IV* polarity phenotypes is notable. These mutants show no polarity defects during cellularization when epithelial cells first form or during gastrulation when the apical junctional belt is assembled. Polarity defects are first seen at stage 11 or 12, and are most prominent at embryonic stage 13. In contrast to wild type, Crb and other apical markers are found throughout the plasma membrane and colocalize with basolateral markers such as Discs large (Dlg, also known as Dlg1) and Fasciclin 3 in *yrt*$^{M/Z}$, *yrt Atpα* and *yrt Nrx-IV* mutants (Fig. 1b, e, g and Supplementary Fig. 2). During late embryogenesis, apical markers become restricted again to the apical membrane, which, however, remains abnormally extended and dome-shaped (Fig. 1b, e, g; stage 17). Thus, Yrt, Na^+,K^+-ATPase and Nrx-IV cooperate and have critical functions in maintaining epithelial polarity during early organogenesis (stages 11–13), well before septate junction assembly is observed[16], and Yrt is recruited to the apical membrane at stage 14 (ref. 2).

Na^+,K^+-ATPase and Nrx-IV are required for septate junction assembly[4,5,17]. Because Yrt also accumulates at septate junctions, we asked whether Yrt has a role in septate junction formation or function. Septate junctions are basal to the adherens junction and, like vertebrate tight junctions, prevent diffusion of solutes between cells[5,13,18]. Dye injection assays show that paracellular barriers in epithelia of *yrt*$^{M/Z}$ embryos are compromised (Supplementary Fig. 3). However, *yrt*$^{M/Z}$ embryos showed a normal complement of septa when examined ultrastructurally (Supplementary Fig. 3) whereas *Atpα* and *Nrx-IV* mutants lack septa[4,17]. Immunoprecipitation experiments demonstrated FERM-domain-dependent interactions between Yrt and the septate junction transmembrane protein Neuroglian (Nrg), and the enrichment of Yrt at septate junctions was less pronounced in *Nrg* mutants (Supplementary Fig. 4).

[1]Department of Cell and Systems Biology, University of Toronto, Toronto, Ontario M5S 3G5, Canada. [2]Department of Medical Biophysics, University of Toronto and The Arthur and Sonia Labatt Brain Tumour Research Centre, Hospital for Sick Children, Toronto, Ontario M5G 1X8, Canada. [3]Department of Biochemistry, Molecular Biology and Cell Biology, Northwestern University, Evanston, Illinois 60208, USA. †Present addresses: Laval University Cancer Research Center, CRCHUQ-Hôtel-Dieu de Québec, Québec City, Québec G1R 2J6,Canada (P.L.); Howard Hughes Medical Institute, Laboratory of Mammalian Cell Biology and Development, The Rockefeller University, New York, New York 10021, USA (S.B.).

However, *Nrg* or *Nrg yrt* mutant embryos did not display overt defects in epithelial polarity. These findings indicate that Yrt is a bona fide septate junction component that has a function distinct from Nrx-IV and Na^+,K^+-ATPase because it is not required for septa formation but essential for barrier function.

Next, we extended our analysis to cytoplasmic adaptor proteins associated with septate junctions. We previously reported that *yrt* does not show genetic interactions with *dlg* or *lgl* (ref. 2)—genes encoding conserved basolateral polarity proteins required for septate junction formation and apical/basal polarity[6,7,19]. We examined *varicose* (*vari*), which encodes a membrane-associated guanylate kinase[20], and *cora*, which encodes a FERM protein that is the *Drosophila* orthologue of mammalian erythrocyte protein band 4.1 and its paralogues (ref. 3). *vari* and *cora* null mutant embryos did not show defects in apical basal polarity. No functional interactions were seen between *yrt* and *vari*. In contrast, *yrt cora* double-mutant embryos displayed marked apicalization defects, with strongly expanded apical membranes and reduced basolateral membranes in all ectodermal epithelia including the epidermis (Fig. 2b, f–i). This phenotype is significantly stronger than the $yrt^{M/Z}$ or *cora* null mutant phenotypes indicating that Yrt and Cora have redundant functions in promoting epithelial polarity.

Similar to $yrt^{M/Z}$ mutants, *yrt cora* double mutants did not show polarity defects during gastrulation. By stage 11, polarity defects were more prominent in *yrt cora* mutants (Fig. 2f, h) than in $yrt^{M/Z}$ embryos (Fig. 1b), which correlates with Cora being expressed from stage 11 onwards (Supplementary Fig. 1). In *yrt cora* mutants, the most severe polarity defects characterized by overlapping distributions of apical and basolateral markers occurred during stage 13 (Fig. 2f, h). However, by late embryogenesis, apical and basolateral markers were segregated (Fig. 2h), although expanded apical membranes and severe defects in tissue organization persisted. Interestingly, the *yrt cora* apicalization phenotype is very similar to

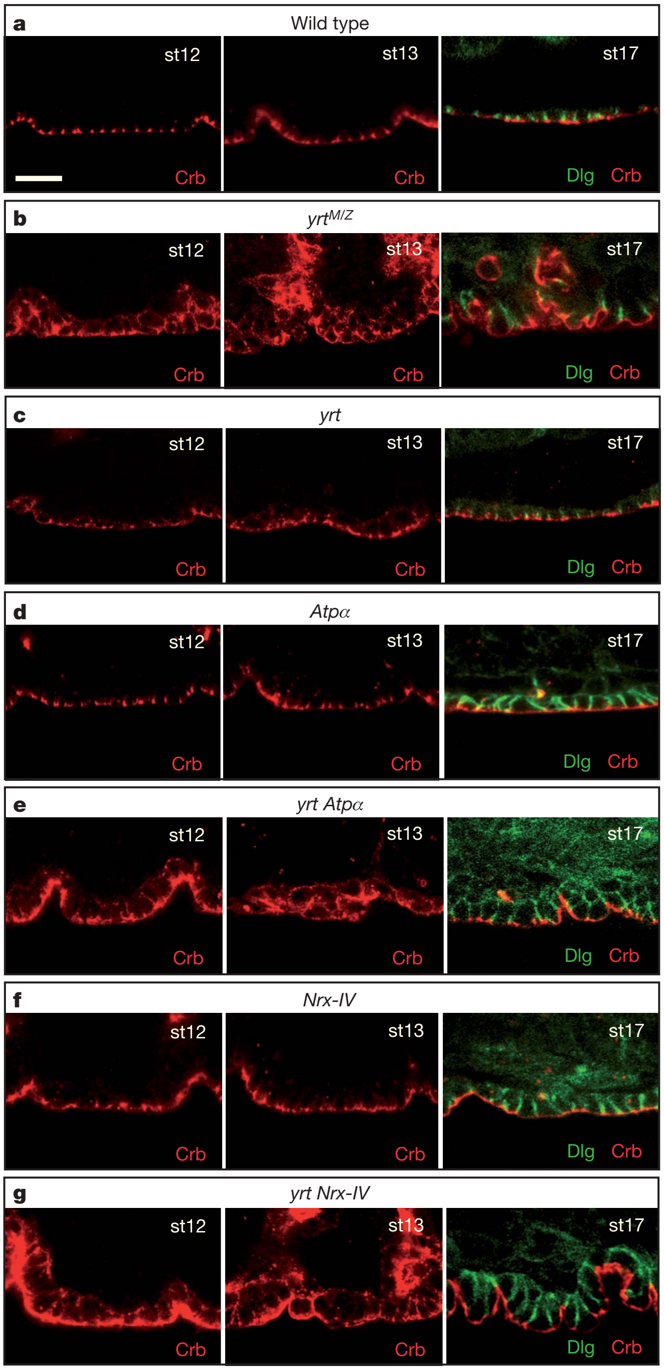


Figure 1 | Synergistic interactions between Yrt, Na^+,K^+-ATPase and Nrx-IV. **a**–**g**, Distribution of Crb and Dlg in wild-type embryos (**a**) and in $yrt^{M/Z}$ (**b**), *yrt* (**c**), *Atpα* (**d**), *yrt Atpα* (**e**), *Nrx-IV* (**f**) and *yrt Nrx-IV* (**g**) mutant embryos at stages (st) 12, 13 and 17. Dlg is only shown in stage-17 embryos. Scale bar, 15 μm.

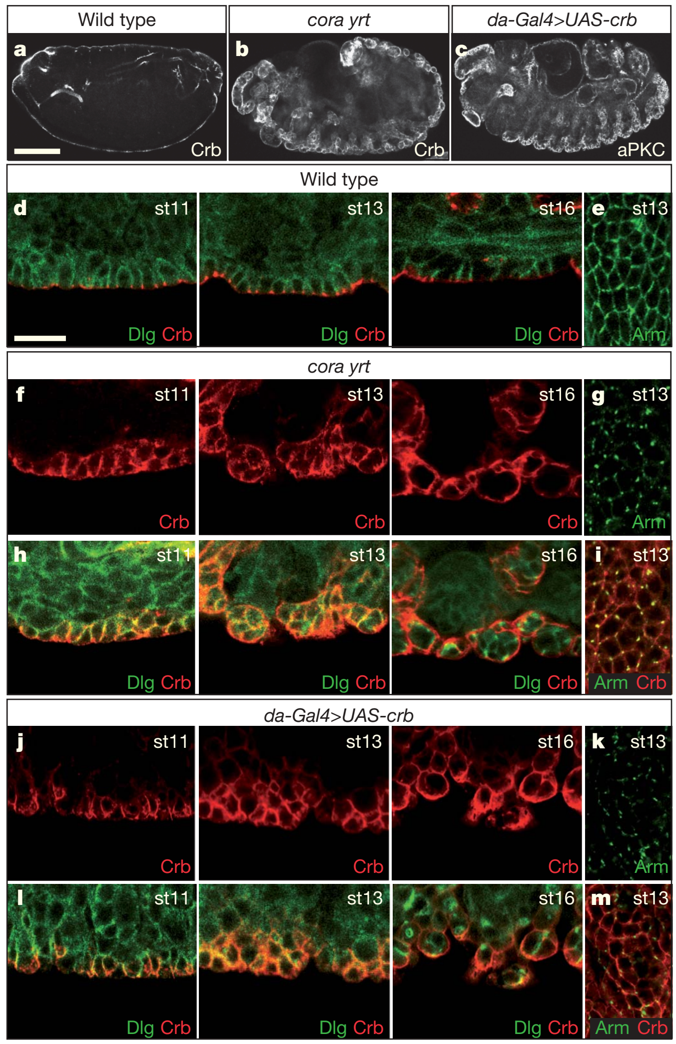


Figure 2 | Excessive apicalization in *cora yrt* double mutants and embryos overexpressing Crb. **a**, **d**, **e**, Wild-type embryos were labelled for the apical marker Crb (**a**), the basolateral marker Dlg and Crb (**d**) or the adherens junction marker Armadillo (Arm; **e**). **b**, **f**–**i**, *cora yrt* double-mutant embryos were labelled for Crb (**b**), Crb and Dlg (**f**, **h**) or Arm and Crb (**g**, **i**). **c**, **j**–**m**, Embryos overexpressing Crb under the control of *da-Gal4* were labelled for the apical marker aPKC (**c**), Crb and Dlg (**j**, **l**) or Arm and Crb (**k**, **m**). Scale bars: **a**–**c**, 100 μm; **d**–**m**, 20 μm.

the phenotype that results from high-level overexpression of Crb (ref. 12) in both timeline and severity (Fig. 2j–m and Supplementary Fig. 5 and 7). The segregation of apical and basolateral markers in late-stage *yrt cora* mutants is also observed in late-stage $yrt^{M/Z}$, *yrt Atpα* and *yrt Nrx-IV* embryos (Fig. 1b, e, g). Thus, polarization mechanisms are at play in late embryos that are not dependent on Yrt and Cora. To address the possibility that this repolarization is the result of redundancy between Yrt/Cora and Lgl-group proteins, which include Scribble (Scrib) as well as Lgl and Dlg, we examined embryos lacking Yrt and Scrib ($yrt^{M/Z}$ $scrib^{M/Z}$). Apical and basolateral markers segregated in $yrt^{M/Z}$ $scrib^{M/Z}$ embryos (Supplementary Fig. 6) indicating that a basolateral polarity mechanism exists in late embryos that is independent of Yrt/Cora and Lgl-group proteins.

Together, our findings suggest that Yrt, Cora, Nrx-IV and the Na^+,K^+-ATPase are a new functionally cooperating group of basolateral epithelial polarity proteins, which we refer to as the 'Yrt/Cora group'. In contrast to the enhancement of the *cora* and *Nrx-IV* null phenotypes by *yrt* mutations, no phenotypic enhancement was seen in *cora Nrx-IV* or *cora Atpα* double mutants, indicating that the Yrt/Cora group is composed of two functionally overlapping pathways. Cora, Nrx-IV and the Na^+,K^+-ATPase belong to one pathway, which is consistent with previously documented biochemical and genetic interactions between these proteins[5,17], whereas Yrt defines a second redundant pathway.

One critical aspect of polarity regulation is that apical and basolateral proteins act antagonistically to set up mutually exclusive membrane domains[6,7,21–23]. In gastrulating *Drosophila* embryos, this interaction occurs between apical factors and basolateral Lgl-group proteins. Similar to *lgl*, *dlg* and *scrib* mutations[6,7], *yrt* mutations partially suppress the *crb* mutant phenotype[2]. Because Yrt can bind Crb directly, we suggested that Yrt negatively regulates Crb function as a component of the apical Crb complex in both apical basal polarity and the control of apical membrane size[2]. The data reported here argue that Yrt opposes Crb function also as a basolateral polarity protein. To test this hypothesis further we asked whether the loss of other Yrt/Cora-group genes could suppress the *crb* mutant phenotype. We found that mutations in *Atpα*, *Nrx-IV* and *cora* ameliorate the epithelial defects of *crb* mutants to an extent similar to that of *yrt* mutations (Fig. 3a–e and Supplementary Fig. 2e). Remarkably, *yrt cora crb* triple-mutant embryos not only showed a suppression of the *crb* mutant phenotype, which was much stronger than the one observed in *yrt crb* or *cora crb* double mutants (Fig. 3c, f and Supplementary Fig. 2e), but also caused a strong suppression of the apicalization effect observed in *yrt cora* mutants (compare Fig. 3f with Fig. 2b, f–i and Supplementary Fig. 7a). These findings further support the conclusion that Yrt and Cora act redundantly, and indicate that mutual competition of basolateral Yrt/Cora-group proteins and apical Crb organizes epithelial membrane domains.

Epithelial differentiation during *Drosophila* embryogenesis can be subdivided into four phases that are characterized by distinct mechanisms governing epithelial polarity (Fig. 3g). (1) Initial cues for polarity are given before or during cellularization[13,24,25]. (2) Fully polarized epithelial cells are established during gastrulation through the interplay of apical Par and Crb complexes and basolateral Lgl-group proteins[6,7]. (3) The Yrt/Cora group acts during organogenesis to promote basolateral polarity and counteracts apical determinants, thereby functionally replacing the Lgl group. (4) The function of Yrt/

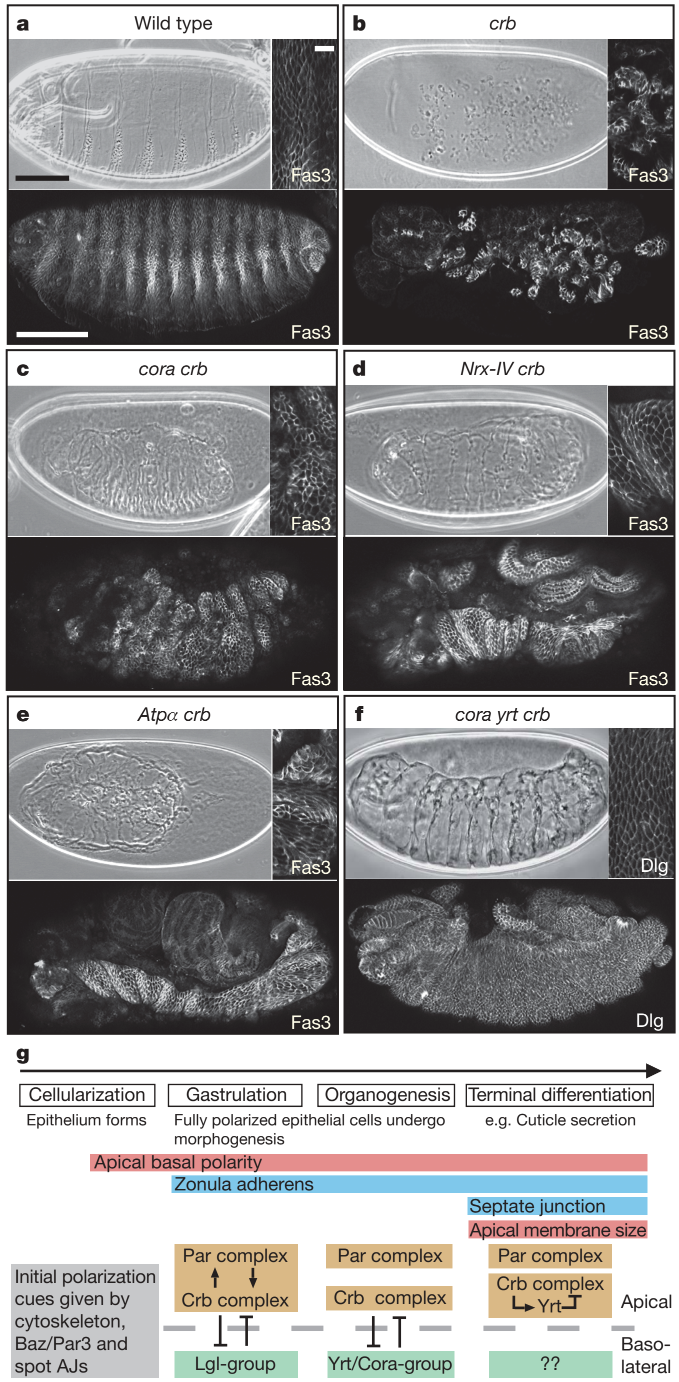


Figure 3 | Mutations in Yrt/Cora-group genes suppress the *crb* mutant phenotype. a–f, Panels show larval cuticles (top left) and Fasciclin 3 (Fas3, **a–e**) or Dlg (**f**) stainings (bottom) of late-stage whole embryos or a close-up of epidermal cells of the same Fas3- or Dlg-labelled embryos (top right) for the following genotypes: wild type (**a**), *crb* (**b**), *cora crb* (**c**), *Nrx-IV crb* (**d**), *Atpα crb* (**e**) and *cora yrt crb* (**f**). The *crb* mutant embryo displays small patches or cysts of epithelial cells. All double mutants (**c–e**) show a substantial patch of continuous epidermal epithelium similar to *crb yrt* double mutants (Supplementary Fig. 2e). The *cora yrt crb* triple-mutant embryo (**f**) shows epidermal epithelium covering almost the entire embryo. Scale bars: cuticle and whole embryos, 100 μm; insets, 20 μm. **g**, Model of regulatory interactions between epithelial polarity proteins during *Drosophila* embryogenesis. Cellularization (stage 5): initial cues for apical basal polarity are given by the cytoskeleton, Baz/Par3 and spot adherens junctions (AJs). Gastrulation (stages 6–10): the apical Par and Crb polarity complexes assemble and are co-dependent. Apical polarity proteins and the basolateral Lgl-group show antagonistic regulatory interactions that set up apical and basal membrane domains, which are segregated by an apical junctional complex, the zonula adherens that forms from spot adherens junctions. Organogenesis (stages 11–13): Par and Crb complexes become independent from each other and act redundantly. Apical polarity proteins and the Yrt/Cora group show antagonistic regulatory interactions that maintain epithelial polarity. Terminal differentiation (stages 14–17): septate junctions form. Yrt is recruited into the apical Crb complex to limit Crb activity in apical membrane size regulation. Unknown factors stabilize the basolateral membrane domain.

Cora and Lgl groups does not account for the polarization of epithelial cells in late embryos, because normalization of polarity is observed in the absence of these factors. This implies the existence of a yet unknown mechanism that stabilizes basolateral polarity. Septate junctions form well after Lgl-group and Yrt/Cora-group proteins have contributed to epithelial polarity, indicating that polarity and septate junction formation are independent functions of these proteins. The identification of a novel group of polarity factors that acts in a discrete time window during development highlights the temporal complexity of the regulation of the epithelial phenotype and further explains why the loss of individual polarity proteins does not completely compromise polarity.

The vertebrate Yrt orthologues Mosaic eyes (Moe, also known as Epb41l5) in zebrafish and EPB41L5 (also known as YMO1 and Limulus) and EPB41L4B (also known as EHM2) in mammals bind to Crb proteins and contribute to epithelial organization[2,8–10,26]. However, as in *Drosophila*, vertebrate Yrt homologues are predominantly associated with the basolateral membrane (Fig. 4a) and seem to be recruited to the apical membrane only at later stages of epithelial development[2,8–11]. To test whether EPB41L5 is required for basolateral differentiation, we used RNA interference in MDCK cells (Fig. 4 and Supplementary Fig. 8). Transient depletion of EPB41L5 using shRNA resulted in a notable cell flattening and an expansion of the cell perimeter, indicating that lateral membranes were strongly reduced and apical and basal membranes were enlarged (Fig. 4a–d and Supplementary Fig. 8g). Consistently, we found that the lateral markers Na^+,K^+-ATPase, Scrib (Fig. 4c) and E-cadherin (not shown) were reduced or lost from the plasma membrane. The same effects were observed using siRNA oligonucleotides targeted to a distinct region of *EPB41L5* (Supplementary Fig. 8a–c). We also established MDCK cells stably transfected with *EPB41L5* short hairpin RNA (shRNA) and examined lateral membrane formation after Ca^{2+}-switch (Fig. 4e and Supplementary Fig. 8d–f, h). After re-addition of Ca^{2+}, *EPB41L5* shRNA cells displayed significant delays in lateral membrane formation and recruitment of E-cadherin to cell–cell contacts. At 8 h after Ca^{2+}-switch, E-cadherin levels appeared similar in control and knockdown cells although experimental cells remained slightly flatter. Interestingly, by 24 h after Ca^{2+}-switch, E-cadherin levels at the plasma membrane had decreased again and appeared lower than in control cells but similar to *EPB41L5* shRNA cells before Ca^{2+}-switch (Fig. 4e and data not shown). *EPB41L5* knockdown cells ultimately established normal cell shape. Whether the formation of a normal lateral membrane in *EPB41L5* shRNA cells is due to residual EPB41L5 expression, or reflects a transient requirement of EPB41L5 for cell polarization as in *Drosophila* epithelia, remains unclear. We conclude that the function of Yrt as a basolateral polarity protein is conserved in mammalian epithelial cells.

The loss of basolateral membrane in *Drosophila* embryos that lack Yrt/Cora-group function and in MDCK cells depleted of EPB41L5 is reminiscent of the loss of basolateral membrane in MDCK cells depleted of the phosphoinositide PtdIns(3,4,5)P_3 (ref. 27) and of human bronchial epithelial cells depleted of β2-spectrin or ankyrin-G[28]. Cora/band 4.1 and the Na^+,K^+-ATPase can associate with either spectrin or the spectrin adaptor ankyrin[29,30], and loss or ectopic expression of EPB41L5 can cause defects in basolateral actin[8]. This raises the possibility that Yrt/Cora-group proteins act by stabilizing the lateral actin/spectrin membrane cytoskeleton. Analysis of vertebrate development in the absence of the function of the Yrt orthologues Moe in zebrafish[11,26] and EPB41L5 in mice[8,10] revealed defects in epithelial organization that reflect Crb-dependent and probably also Crb-independent functions of Yrt proteins. For example, the abnormal cell shape and multi-layering defects seen in the developing neuroepithelium of mutant mouse embryos could result from defects in basolateral polarity. Moreover, mouse embryos lacking EPB41L5 show defects in epithelial–mesenchymal transition of mesodermal cells during gastrulation[8,10]. These surprising findings indicate that Yrt proteins are core regulators of animal tissue organization that can enhance epithelial or mesenchymal cell differentiation.

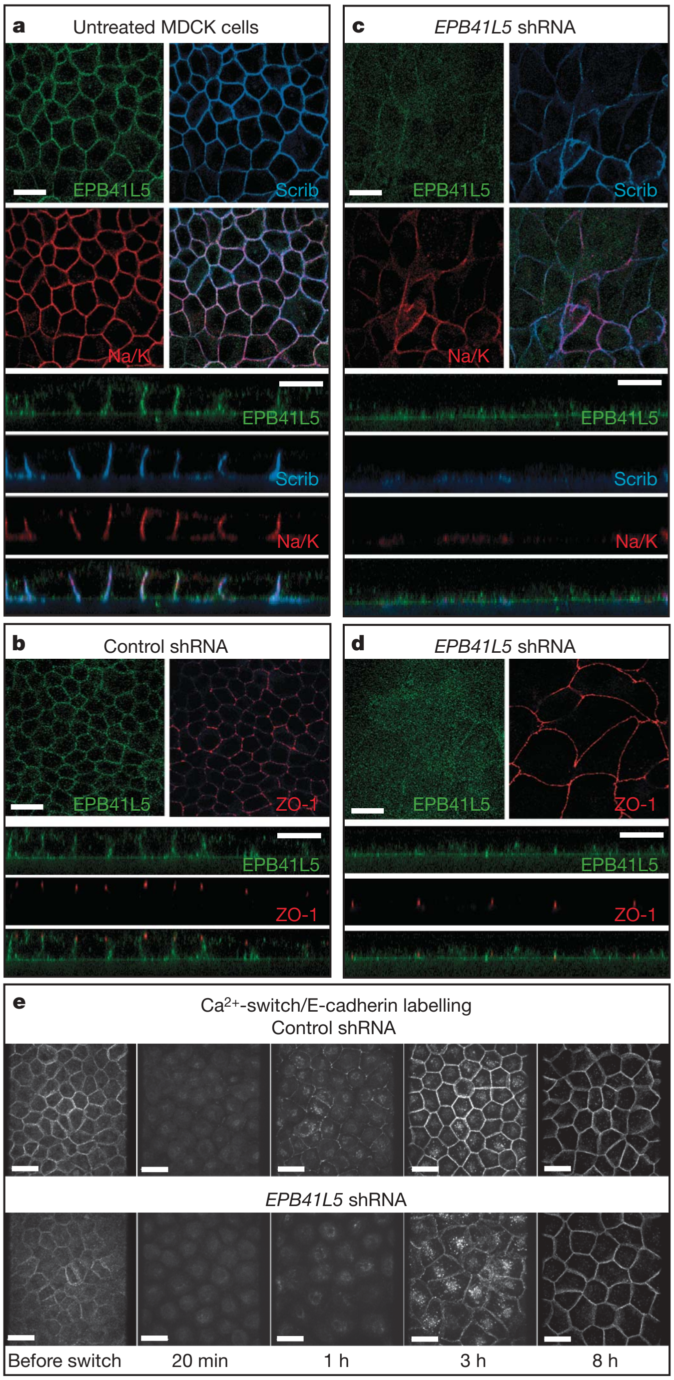


Figure 4 | Knockdown of EPB41L5 in MDCK cells compromises lateral membrane formation. **a**, Polarized MDCK cells were stained for endogenous EPB41L5, Na^+,K^+-ATPase α-subunit (Na/K) and Scrib. EPB41L5 colocalizes with the Na^+,K^+-ATPase and Scrib at the lateral membrane. **b–d**, MDCK cells were transfected with control luciferase shRNA (**b**) or *EPB41L5* shRNA (**c**, **d**) and stained for EPB41L5, Na^+,K^+-ATPase, Scrib or ZO-1 as indicated. *EPB41L5*-shRNA-treated cells display strongly reduced lateral membranes, enlarged apical membranes and reduced levels of Na^+,K^+-ATPase and Scrib. **e**, Ca^{2+}-switch experiment in MDCK cells with stable EPB41L5 knockdown. Control (luciferase)-shRNA or *EPB41L5*-shRNA-expressing MDCK cells were depleted of Ca^{2+} overnight, which caused a loss of cell–cell contact and lateral membrane. After addition of Ca^{2+} for the indicated time period, cells were labelled for E-cadherin. Scale bars, 10 μm.

METHODS SUMMARY

***Drosophila* genetics.** The analysis presented here is based on the evaluation of 48 different single-, double- or triple-mutant genotypes using immunocytochemistry, electron microscopy and preparations of the larval cuticle. To identify Yrt binding partners through immunoprecipitation we used the Yrt antibody GP7 (ref. 2) or generated transgenic animals expressing Flag-tagged Yrt isoforms under Gal4/UAS control. The Yrt isoforms were full-length Yrt-β (ref. 2), Yrt-β lacking the carboxy-terminal PDZ domain binding site (Yrt-$\beta^{\Delta PDZ}$), and the FERM domain of Yrt (Yrt^{FERM}).

***EPB41L5* knockdown in MDCK cells.** Transient knockdown of *EPB41L5* was accomplished with shRNA (5′-GGAGCTAACCCGGTATTTATT-3′; using luciferase shRNA as control) or short interfering RNA (siRNA) (5′-CGACTAGGATCACGATTTA-3′; using scrambled siRNA as control) directed against a different region of the *EPB41L5* transcript. Stable *EPB41L5*-knockdown MDCK cell lines were generated by transfecting cells with shRNA expression constructs followed by positive selection. MDCK cells were analysed by immunocytochemistry.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 13 December 2008; accepted 5 April 2009.

1. Nelson, W. J. Adaptation of core mechanisms to generate cell polarity. *Nature* **422,** 766–774 (2003).
2. Laprise, P. *et al.* The FERM protein Yurt is a negative regulatory component of the Crumbs complex that controls epithelial polarity and apical membrane size. *Dev. Cell* **11,** 363–374 (2006).
3. Fehon, R. G., Dawson, I. A. & Artavanis-Tsakonas, S. A *Drosophila* homologue of membrane-skeleton protein 4.1 is associated with septate junctions and is encoded by the coracle gene. *Development* **120,** 545–557 (1994).
4. Baumgartner, S. *et al.* A *Drosophila* neurexin is required for septate junction and blood-nerve barrier formation and function. *Cell* **87,** 1059–1068 (1996).
5. Paul, S. M., Ternet, M., Salvaterra, P. M. & Beitel, G. J. The Na^+/K^+ ATPase is required for septate junction function and epithelial tube-size control in the *Drosophila* tracheal system. *Development* **130,** 4963–4974 (2003).
6. Tanentzapf, G. & Tepass, U. Interactions between the *crumbs*, *lethal giant larvae* and *bazooka* pathways in epithelial polarization. *Nature Cell Biol.* **5,** 46–52 (2003).
7. Bilder, D., Schober, M. & Perrimon, N. Integrated activity of PDZ protein complexes regulates epithelial polarity. *Nature Cell Biol.* **5,** 53–58 (2003).
8. Lee, J. D., Silva-Gagliardi, N. F., Tepass, U., McGlade, C. J. & Anderson, K. V. The FERM protein Epb4.1l5 is required for organization of the neural plate and for the epithelial–mesenchymal transition at the primitive streak of the mouse embryo. *Development* **134,** 2007–2016 (2007).
9. Gosens, I. *et al.* FERM protein EPB41L5 is a novel member of the mammalian CRB–MPP5 polarity complex. *Exp. Cell Res.* **313,** 3959–3970 (2007).
10. Hirano, M., Hashimoto, S., Yonemura, S., Sabe, H. & Aizawa, S. EPB41L5 functions to post-transcriptionally regulate cadherin and integrin during epithelial–mesenchymal transition. *J. Cell Biol.* **182,** 1217–1230 (2008).
11. Hsu, Y. C., Willoughby, J. J., Christensen, A. K. & Jensen, A. M. Mosaic Eyes is a novel component of the Crumbs complex and negatively regulates photoreceptor apical size. *Development* **133,** 4849–4859 (2006).
12. Wodarz, A., Hinz, U., Engelbert, M. & Knust, E. Expression of crumbs confers apical character on plasma membrane domains of ectodermal epithelia of *Drosophila. Cell* **82,** 67–76 (1995).
13. Tepass, U., Tanentzapf, G., Ward, R. & Fehon, R. Epithelial cell polarity and cell junctions in *Drosophila. Annu. Rev. Genet.* **35,** 747–784 (2001).
14. Pellikka, M. *et al.* Crumbs, the *Drosophila* homologue of human CRB1/RP12, is essential for photoreceptor morphogenesis. *Nature* **416,** 143–149 (2002).
15. Wu, V. M. & Beitel, G. J. A junctional problem of apical proportions: epithelial tube-size control by septate junctions in the *Drosophila* tracheal system. *Curr. Opin. Cell Biol.* **16,** 493–499 (2004).
16. Tepass, U. & Hartenstein, V. The development of cellular junctions in the *Drosophila* embryo. *Dev. Biol.* **161,** 563–596 (1994).
17. Genova, J. L. & Fehon, R. G. Neuroglian, Gliotactin, and the Na^+/K^+ ATPase are essential for septate junction function in *Drosophila. J. Cell Biol.* **161,** 979–989 (2003).
18. Lamb, R. S., Ward, R. E., Schweizer, L. & Fehon, R. G. *Drosophila coracle*, a member of the protein 4.1 superfamily, has essential structural functions in the septate junctions and developmental functions in embryonic and adult epithelial cells. *Mol. Biol. Cell* **9,** 3505–3519 (1998).
19. Bilder, D., Li, M. & Perrimon, N. Cooperative regulation of cell polarity and growth by *Drosophila* tumor suppressors. *Science* **289,** 113–116 (2000).
20. Wu, V. M. *et al. Drosophila* Varicose, a member of a new subgroup of basolateral MAGUKs, is required for septate junctions and tracheal morphogenesis. *Development* **134,** 999–1009 (2007).
21. Benton, R. & St Johnston, D. *Drosophila* PAR-1 and 14–3-3 inhibit Bazooka/PAR-3 to establish complementary cortical domains in polarized cells. *Cell* **115,** 691–704 (2003).
22. Yamanaka, T. *et al.* Mammalian Lgl forms a protein complex with PAR-6 and aPKC independently of PAR-3 to regulate epithelial cell polarity. *Curr. Biol.* **13,** 734–743 (2003).
23. Suzuki, A. *et al.* aPKC acts upstream of PAR-1b in both the establishment and maintenance of mammalian epithelial polarity. *Curr. Biol.* **14,** 1425–1435 (2004).
24. Harris, T. J. & Peifer, M. Adherens junction-dependent and -independent steps in the establishment of epithelial cell polarity in *Drosophila. J. Cell Biol.* **167,** 135–147 (2004).
25. Mavrakis, M., Rikhy, R. & Lippincott-Schwartz, J. Plasma membrane polarity and compartmentalization are established before cellularization in the fly embryo. *Dev. Cell* **16,** 93–104 (2009).
26. Jensen, A. M. & Westerfield, M. Zebrafish mosaic eyes is a novel FERM protein required for retinal lamination and retinal pigmented epithelial tight junction formation. *Curr. Biol.* **14,** 711–717 (2004).
27. Gassama-Diagne, A. *et al.* Phosphatidylinositol-3,4,5-trisphosphate regulates the formation of the basolateral plasma membrane in epithelial cells. *Nature Cell Biol.* **8,** 963–970 (2006).
28. Kizhatil, K. *et al.* Ankyrin-G and 2-spectrin collaborate in biogenesis of lateral membrane of human bronchial epithelial cells. *J. Biol. Chem.* **282,** 2029–2037 (2007).
29. Hortsch, M. & Margolis, B. Septate and paranodal junctions: kissing cousins. *Trends Cell Biol.* **13,** 557–561 (2003).
30. Bennett, V. & Healy, J. Organizing the fluid membrane bilayer: diseases linked to spectrin and ankyrin. *Trends Mol. Med.* **14,** 28–36 (2008).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We would like to thank V. Auld, J. Casanova, H. Bellen, M. Hortsch, R. Fehon, M. Bhat, R. Schuh, the Developmental Studies Hybridoma Bank and the Bloomington *Drosophila* Stock Center for reagents. We thank D. Godt, R. Winklbauer and T. Harris for critical comments on the manuscript, and H. Hong, M. Pellikka and W. Russin for technical assistance. This work was supported by postdoctoral fellowships from the CIHR (to P.L.), a predoctoral fellowship from the Vision Science Research Program, University of Toronto (to S.B.) and by NIH Lung Biology Training Grant 5 (to S.M.P.). Operating support was provided by the CIHR (to U.T. and P.L.), the Foundation Fighting Blindness Canada (to C.J.M.) and the NIH (to G.J.B.).

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to U.T. (u.tepass@utoronto.ca).

METHODS

***Drosophila* genetics.** The alleles used in this study were: yrt^{75a} and yrt^{65a} (ref. 2), yrt^{E99} (ref. 31), $cora^{5}$ (ref. 18), $cora^{1}$ (ref. 3), crb^{11A22} (ref. 32), $Atp\alpha^{DTS1R2}$ (ref. 33), $nrv2^{23B}$ (ref. 17), $Nrx\text{-}IV^{4865}$ (ref. 4), Nrg^{17} (ref. 34), $Gli^{AE2\Delta4b}$ (ref. 35), $Fas2^{EB112}$ (ref. 36), $Fas3^{E25}$ (ref. 37), $mega^{EA97}$ (ref. 38), $sinu^{06524}$ (ref. 38), $Cont^{ex956}$ (ref. 39), $Lac^{BG01462}$ (ref. 40). Recombinant chromosomes generated were: $Nrx\text{-}IV^{4865}\,yrt^{75a}$, $yrt^{75a}\,Atp\alpha^{DTS1R2}$, $Nrx\text{-}IV^{4865}\,crb^{11A22}$, $crb^{11A22}\,Atp\alpha^{DTS1R2}$, $sinu^{06524}\,yrt^{75a}$, $cont^{ex956}\,yrt^{75a}$, FRT82B $yrt^{75a}\,scrib^{1}$,$Gli^{AEZ\Delta4b}\,cora^{5}$, $Fas3^{E25}\,cora^{5}$, $Lac^{BG01462}\,cora^{5}$, $sinu^{06524}\,yrt^{75a}$, $Cont^{ex956}\,yrt^{75a}$. Genetic interactions were observed in the following allelic combinations: $Nrx\text{-}IV^{4865}\,yrt^{75a}$, $yrt^{75a}\,Atp\alpha^{DTS1R2}$, $Nrx\text{-}IV^{4865}\,crb^{11A22}$, $crb^{11A22}\,Atp\alpha^{DTS1R2}$, $cora^{5}\,yrt^{75a}$, $cora^{1}\,yrt^{E99}$, $cora^{1}\,yrt^{75a}$, $cora^{5}\,yrt^{E99}$, $cora^{5}\,crb^{11A22}$, $crb^{11A22}\,yrt^{75a}$, $cora^{5}\,yrt^{75a}\,crb^{11A22}$. No obvious genetic interactions were observed in the following double-mutant combinations: $nrg^{17}\,crb^{11A22}$, $nrg^{17}\,yrt^{75a}$, $Gli^{AE2\Delta4b}\,yrt^{75a}$, $Fas3^{E25}\,yrt^{75a}$, $mega^{EA97}\,yrt^{75a}$, $Lac^{BG01462}\,yrt^{75a}$, $sinu^{06524}\,yrt^{75a}$, $cont^{ex956}\,yrt^{75a}$, $cora^{5}\,Nrx\text{-}IV^{4865}$, $nrg^{17}\,cora^{5}$, $Gli^{AE2\Delta4b}\,cora^{5}$, $Fas2^{EB112}\,cora^{5}$, $Fas3^{E25}\,cora^{5}$, $mega^{EA97}\,cora^{5}$, $Lac^{BG01462}\,cora^{5}$, $cora^{5}\,Cont^{ex956}$, $cora^{5}\,sinu^{06524}$, $cora^{5}\,Atp\alpha^{DTS1R2}$, $sinu^{nw7}\,crb^{11A22}$. Germline clones were generated using the FLP-DFS (FLP-recombinase-dominant female sterile) technique (ref. 41). Overexpression of Crb or Yrt was accomplished by crossing $UAS\text{–}crb^{wt2e}$ (ref. 12), $UAS\text{–}yrt\text{-}\beta$, $UAS\text{–}yrt^{FERM}$ or $UAS\text{–}yrt\text{-}\beta^{\Delta PDB}$ to the ubiquitous driver *da–GAL4* (ref. 12).

***Drosophila* immunocytochemistry.** *Drosophila* embryos were fixed as previously described (ref. 32) except for Arm, where embryos were heat-fixed (ref. 42). Primary antibodies used were: guinea pig anti-Yrt GP7 (ref. 2), rabbit anti-Nrv2 (ref. 43), rat anti-Vari (ref. 20), rabbit anti-aPKC C-20 (Santa Cruz Biotechnology), mouse monoclonal anti-Arm, anti-Cora, anti-ATPase α (a5), anti-Fas3 and anti-Dlg (Developmental Studies Hybridoma Bank) and rat anti-Crb (ref. 14). Secondary antibodies were conjugated to Cy3, Cy5 (Jackson Immunoresearch Laboratories) or Alexafluor 488 (Molecular Probes). Confocal images were collected on a Zeiss LSM 510 confocal microscope with ×40 or ×63 oil objectives. Images were processed and arranged in Adobe Photoshop and Adobe Illustrator.

Plasmid constructs and transgenic lines. Full-length cDNA encoding Yrt-β (ref. 2) was subcloned into the XhoI site of pBS-SK from overlapping expressed sequence tags LD01167 and RE35248 (Berkeley *Drosophila* Genome Project). Yrt-β, Yrt-β with the C-terminal PDZ domain binding (PDB) site deleted (Yrt-$\beta^{\Delta PDB}$; amino acids 1–968) and the FERM domain of Yrt (Yrt^{FERM}; amino acids 39–272) were subcloned into pENTRC (Invitrogen) and transferred using Gateway Technology (Invitrogen) into pTFW-attB, an expression vector that adds a 3×Flag tag to the amino terminus. pTFW-attB was produced by adding an *attB* site[44] into the NsiI site of pTFW (*Drosophila* Gateway Vector Collection; http://www.ciwemb.edu/labs/murphy/Gateway%20vectors.html). Transgenic animals were produced by Genetic Services, Inc., using flies carrying *attP2* (ref. 44).

Electron microscopy. Embryos were prepared for transmission electron microscopy as described[7] except that post-fixation was carried out in 1% OsO_4, 0.8% w/v potassium ferricyanide in 100 mM cacodylate buffer on ice and embryos were embedded in Spurr's resin.

Septate junction permeability assay. Transepithelial permeability assays were carried out as described[5]. Stage-17 embryos were injected with 10 kDa Texas-Red dextran, and dye diffusion across epithelia of the main trachea trunk was monitored by confocal microscopy 10 min after injection. Of six injected wild-type embryos, none showed leakage of dye into the tracheal lumen. All three injected $Atp\alpha^{DTS1R2}$ mutants and more than ten $nrv2^{23B}$ mutants showed strong leakage. Four injected yrt^{E99} mutant embryos showed consistent weak leakage of dye into the tracheal lumen. Nineteen embryos derived from females carrying a yrt^{E99} mutant germline and crossed to $yrt^{E99}/+$ males were injected. Fifty per cent of embryos are expected to be $yrt^{M/Z}$, and thus *yrt* null, and the other 50% to contain a paternal wild-type copy of *yrt* and develop normally. We found that 11 embryos showed strong leakage of dye into the tracheal lumen. These embryos were positively identified as $yrt^{M/Z}$ embryos on the basis of their morphological defects.

Immunoprecipitations and immunoblotting. *Drosophila* embryos were homogenized in ice-cold 1% Triton X-100 buffer[45] and embryo lysates were cleared by centrifugation. 1 mg of each lysate was incubated with 2 μg of anti-Flag (Sigma) under constant agitation for 2 h at 4 °C. 40 μg of protein G–Sepharose beads were added to lysates and incubated under constant agitation for 2 h at 4 °C. Immunocomplexes were washed five times with lysis buffer and subjected to immunoblot analysis using the following primary antibodies: anti-Yrt GP7 (ref. 2) 1/1,000, anti-Nrx-IV (ref. 4) 1/500, or anti-Nrg 3F4 (ref. 46) 1/1,000. Horseradish peroxidase (HRP)-coupled secondary antibodies (Jackson Immunoresearch Laboratories) were used at a dilution of 1/1,000. Co-immunoprecipitation experiments with Flag-tagged Yrt constructs were negative for the following antibodies: anti-Fasciclin 2 (1D4, DSHB) 1/1,000, anti-Gliotactin (ref. 35) 1/500, anti-Nrx-IV (ref. 4) 1/500, anti-Lachesin (ref. 40) 1/5,000 and anti-Contactin (ref. 39) 1/2,000.

***EPB41L5* knockdown in MDCK cells.** Transient shRNA knockdown of EPB41L5 was carried out by transfecting MDCK cells at 75% confluence with 1 μg of *EPB41L5* shRNA (5′-GGAGCTAACCCGGTATTTATT-3′; Open Biosystems, oligo ID V2MM37774, corresponding to nucleotides 395–416 of mouse *EPB41L5*) or control luciferase shRNA using Lipofectamine 2000 (Invitrogen) according to the manufacturer's directions. Cells were fixed and stained 48 h post-transfection. For transient siRNA knockdown a siRNA oligonucleotide to target canine *EPB41L5* (5′-CGACTAGGATCACGATTTA-3′) was designed using siDESIGN Center (Dharmacon) and purchased from Dharmacon. Transient siRNA oligonucleotide knockdown of EPB41L5 was carried out by transfecting MDCK cells at 75% confluence with 200 pmol siRNA duplex or scrambled RNA using Lipofectamine 2000 (Invitrogen). Cells were fixed and stained 24 h post-transfection. To generate stably transfected cell lines, MDCK cells were cultured in Dulbecco's Modified Eagle Medium (Wisent Inc.) supplemented with 5% fetal bovine serum. Cells were transfected with 2 μg of the *EPB41L5* shRNA in the pSM2c vector (Open Biosystems, oligo ID V2MM37774) or control luciferase shRNA using Lipofectamine 2000 (Invitrogen) according to the manufacturer's instructions to generate control and EPB41L5 stable knockdown lines. After selection with 2 mg ml^{-1} of puromycin for two weeks, clones were isolated and propagated to generate stable cell line clones.

Ca^{2+}-switch experiments. Ca^{2+}-switch experiments were conducted by incubating cells overnight in low calcium medium (SMEM, Gibco) followed by re-addition of normal growth medium (DMEM supplemented with 5% FBS). Cells were fixed at different time points, as indicated, and stained.

MDCK immunocytochemistry. MDCK cells were grown on Falcon cell culture inserts to confluence and fixed with 2% PFA, 30 mM sucrose for 20 min. Cells were permeablized in 0.05% saponin, and blocked in 3% normal donkey serum. Primary antibodies were applied for 1 h at 37 °C. Cells were washed as follows: 1 × 10 min with PBS, 3 × 10 min with 0.05% Triton in PBS and 1 × 10 min with PBS. Secondary antibodies were applied for 30 min at 37 °C; cells were washed as described above, and mounted in Dako Cytomation fluorescent mounting medium. Antibodies used to label MDCK cells were: anti-YMO1/EPB41L5 rabbit polyclonal antibody, raised against the C-terminal 60 amino acids of human EPB41L5 (BC032822), and used at a dilution of 1:50; anti-E-cadherin, 1:500 (Sigma); anti-Na^{+},K^{+}-ATPase α, 1:500 (Abcam); anti-ZO-1, 1:300 (Zymed); anti-Scribble, 1:200 (Santa Cruz). Confocal images were acquired using either a Zeiss LSM 510 microscope at ×100 magnification, and image analysis was performed using LSM software. Alternatively, confocal images were acquired using a Quorum spinning disk confocal microscope at ×63 magnification, and image analysis was performed using Volocity software. Images were processed and arranged in Adobe Photoshop and Adobe Illustrator.

31. Hoover, K. B. & Bryant, P. J. *Drosophila* Yurt is a new protein-4.1-like protein required for epithelial morphogenesis. *Dev. Genes Evol.* **212**, 230–238 (2002).
32. Tepass, U., Theres, C. & Knust, E. *Crumbs* encodes an EGF-like protein expressed on apical membranes of *Drosophila* epithelial cells and required for organization of epithelia. *Cell* **61**, 787–799 (1990).
33. Palladino, M. J., Bower, J. E., Kreber, R. & Ganetzky, B. Neural dysfunction and neurodegeneration in *Drosophila* Na^{+}/K^{+} ATPase alpha subunit mutants. *J. Neurosci.* **23**, 1276–1286 (2003).
34. Bieber, A. J. *et al.* *Drosophila* neuroglian: a member of the immunoglobulin superfamily with extensive homology to the vertebrate neural adhesion molecule L1. *Cell* **59**, 447–460 (1989).
35. Auld, V. J., Fetter, R. D., Broadie, K. & Goodman, C. S. Gliotactin, a novel transmembrane protein on peripheral glia, is required to form the blood-nerve barrier in *Drosophila*. *Cell* **81**, 757–767 (1995).
36. Grenningloh, G., Rehm, E. J. & Goodman, C. S. Genetic analysis of growth cone guidance in drosophila: Fasciclin II functions as a neuronal recognition molecule. *Cell* **67**, 45–57 (1991).
37. Whitlock, K. Development of *Drosophila* wing sensory neurons in mutants with missing or modified cell surface molecules. *Development* **117**, 1251–1260 (1993).
38. Beitel, G. J. & Krasnow, M. A. Genetic control of epithelial tube size in the *Drosophila* tracheal system. *Development* **127**, 3271–3282 (2000).
39. Faivre-Sarrailh, C. *et al.* *Drosophila* contactin, a homolog of vertebrate contactin, is required for septate junction organization and paracellular barrier function. *Development* **131**, 4931–4942 (2004).
40. Llimargas, M., Strigini, M., Katidou, M., Karagogeos, D. & Casanova, J. Lachesin is a component of a septate junction-based mechanism that controls tube size and epithelial integrity in the *Drosophila* tracheal system. *Development* **131**, 181–190 (2004).
41. Chou, T. B. & Perrimon, N. The autosomal Flp–Dfs technique for generating germline mosaics in *Drosophila melanogaster*. *Genetics* **144**, 1673–1679 (1996).
42. Tepass, U. Crumbs, a component of the apical membrane, is required for zonula adherens formation in primary epithelia of *Drosophila*. *Dev. Biol.* **177**, 217–225 (1996).

43. Paul, S. M., Palladino, M. J. & Beitel, G. J. A pump-independent function of the Na,K-ATPase is required for epithelial junction function and tracheal tube-size control. *Development* 134, 147–155 (2007).
44. Groth, A. C., Fish, M., Nusse, R. & Calos, M. P. Construction of transgenic *Drosophila* by using the site-specific integrase from phage C31. *Genetics* 166, 1775–1782 (2004).
45. Laprise, P. *et al.* Phosphatidylinositol 3-kinase controls human intestinal epithelial cell differentiation by promoting adherens junction assembly and p38 MAPK activation. *J. Biol. Chem.* 277, 8226–8234 (2002).
46. Hortsch, M., Wang, Y. M., Marikar, Y. & Bieber, A. J. The cytoplasmic domain of the *Drosophila* cell adhesion molecule neuroglian is not essential for its homophilic adhesive properties in S2 cells. *J. Biol. Chem.* 270, 18809–18817 (1995).

LETTERS

Structural insight into the autoinhibition mechanism of AMP-activated protein kinase

Lei Chen[1]*, Zhi-Hao Jiao[1]*, Li-Sha Zheng[2]*, Yuan-Yuan Zhang[1], Shu-Tao Xie[1], Zhi-Xin Wang[1,2] & Jia-Wei Wu[1]

The AMP-activated protein kinase (AMPK) is characterized by its ability to bind to AMP, which enables it to adjust enzymatic activity by sensing the cellular energy status and maintain the balance between ATP production and consumption in eukaryotic cells[1,2]. It also has important roles in the regulation of cell growth and proliferation, and in the establishment and maintenance of cell polarity[3]. These important functions have rendered AMPK an important drug target for obesity, type 2 diabetes and cancer treatments[4]. However, the regulatory mechanism of AMPK activity by AMP binding remains unsolved. Here we report the crystal structures of an unphosphorylated fragment of the AMPK α-subunit (KD-AID) from *Schizosaccharomyces pombe* that contains both the catalytic kinase domain and an autoinhibitory domain (AID), and of a phosphorylated kinase domain from *Saccharomyces cerevisiae* (Snf1-pKD). The AID binds, from the 'backside', to the hinge region of its kinase domain, forming contacts with both amino-terminal and carboxy-terminal lobes. Structural analyses indicate that AID binding might constrain the mobility of helix αC, hence resulting in an autoinhibited KD-AID with much lower kinase activity than that of the kinase domain alone. AMP activates AMPK both allosterically and by inhibiting dephosphorylation[5,6]. Further *in vitro* kinetic studies demonstrate that disruption of the KD-AID interface reverses the autoinhibition and these AMPK heterotrimeric mutants no longer respond to the change in AMP concentration. The structural and biochemical data have shown the primary mechanism of AMPK autoinhibition and suggest a conformational switch model for AMPK activation by AMP.

AMPKs are highly conserved heterotrimeric enzymes found in most eukaryotic species. The catalytic α-subunit contains a conventional Ser/Thr kinase domain, followed by an autoinhibitory sequence and a C-terminal segment for interacting with the β-subunit[7,8] (Supplementary Fig. 1). The activity of AMPK is tightly regulated by upstream kinases through the phosphorylation of a conserved threonine (Thr 172 in rat) within the activation segment[9]. Furthermore, 5′-AMP that binds to the regulatory γ-subunit activates AMPK by eliciting allosteric changes and inhibiting pThr 172 dephosphorylation[5,6,10]. The scaffold β-subunit bridges α- and γ-subunits by means of its C-terminal sequence, and contains a central non-catalytic glycogen-binding domain that may sense the status of cellular energy reserved in the form of glycogen[11,12]. In addition to AMP, the thienopyridone A-769662 was reported to directly activate AMPK by an unexpected mechanism involving the glycogen-binding domain of the β1-subunit, highlighting a regulatory role of the β-subunit in modulating AMPK activity[13–16]. Recently, the core structures of the αβγ heterotrimers from yeast and mammals, containing the whole γ-subunit but only the C-terminal interacting segments of the α- and β-subunits, have been independently determined[17–19]. Because the critical kinase domain and autoinhibitory sequence were not included, many outstanding questions concerning the regulatory mechanism of AMPK activity remain, such as how the catalytic kinase domain is regulated by the autoinhibitory sequence and how AMP binding to the γ-subunit can ultimately enhance the kinase activity in the α-subunit.

We generated and purified the kinase domains and KD-AID fragments from rat and yeast AMPKs, and all proteins could be readily phosphorylated and activated by CaMKKβ[15] (Supplementary Fig. 2). Like mammalian AMPK, the yeast KD-AID fragments were inactive in their unphosphorylated state, and exhibited low basal kinase activities when phosphorylated at the Thr residues (Thr 189 in *S. pombe* and Thr 210 in *S. cerevisiae*; Fig. 1a–c). The differences between the enzymatic activities of yeast kinase domains and those of KD-AID fragments were found to be around tenfold, comparable to the 17-fold increase of the rat kinase domain to its KD-AID[7]. These data demonstrate that the autoinhibitory property is evolutionarily conserved.

To determine the molecular and structural basis for the autoinhibition by the AID and the activation by upstream kinases, we solved a 2.8-Å structure of the unphosphorylated KD-AID fragment from *S. pombe* (Fig. 1d), and a 2.9-Å structure of the phosphorylated kinase domain from *S. cerevisiae* (Snf1-pKD) (Fig. 1e and Supplementary Fig. 3). Both kinase domains have the same canonical Ser/Thr protein kinase fold as those from *S. cerevisiae* Snf1 (Protein Data Bank (PDB) accession codes 3FAM and 2FH9)[20,21] and human AMPK-α2 (also known as PRKAA2; PDB accession codes 2H6D and 2YZA). The N-terminal lobe consists of a five-stranded antiparallel β-sheet and two α-helices, including the prominent helix αC, whereas the C-terminal lobe is dominated by α-helices. Snf1-pKD displays a 'closed' conformation, although the phosphorylated Thr 210 is invisible, whereas the unphosphorylated kinase domain of *S. pombe* KD-AID adopts an 'open' inactive conformation readily superimposed to Snf1-KD[22] (Fig. 1f). The AID adopts a compact conformation consisting of three α-helices (α1–α3), and binds to the backside of the kinase domain through a new mechanism (see later).

Both structures comprise two molecules within an asymmetric unit that form head-to-tail dimers involving the activation segments and helices αG. Notably, the well-ordered activation segment of *S. pombe* KD-AID is jammed into its active site, and the critical catalytic residues make several intra- and intermolecular contacts (Supplementary Fig. 4). Such a dimeric interaction is expected to inhibit its kinase activity. To determine the biological significance of crystal packing interfaces of dimeric kinase domains[23], we first characterized the oligomeric state of AMPK fragments in solution (Supplementary Fig. 5). Unlike the kinase domain or the KD-AID fragment from *S. cerevisiae* that showed phosphorylation-dependent dimer–monomer switch, both unphosphorylated and phosphorylated *S. pombe* KD-AID were eluted as monomers. Similar results were obtained with the *S. pombe*

[1]MOE Key Laboratory of Bioinformatics, Department of Biological Sciences and Biotechnology, Tsinghua University, Beijing 100084, China. [2]Institute of Biophysics and Graduate University, Chinese Academy of Sciences, Beijing 100101, China.
*These authors contributed equally to this work.

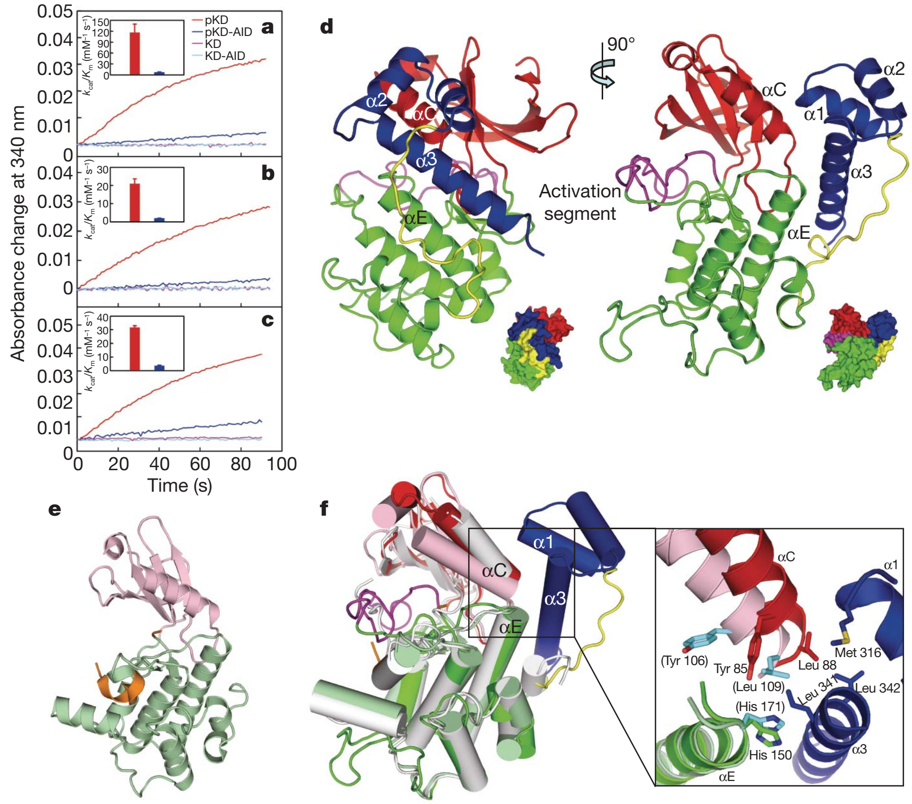


Figure 1 | Structure and activity of the AMPK kinase domain and KD-AID. a–c, Time courses of AMPK-catalysed SAMS (a modified peptide derived from residues 73–85 of rat acetyl-CoA carboxylase) phosphorylation. The absorption was recorded after the addition of 100 nM rat (**a**), 500 nM *S. pombe* (**b**), or 500 nM *S. cerevisiae* (**c**) AMPK proteins. The insets show the $k_{cat}/K_{m(SAMS)}$ (Michaelis constant) values for phosphorylated proteins (mean and s.e.m., $n = 3$). **d**, Schematic and surface representations of *S. pombe* KD-AID in two views related by a 90° rotation around a vertical axis. The N- and C-lobes of the kinase domain are coloured in red and green, respectively, the activation segment is in magenta, the linker is in yellow, and the AID is in blue. **e**, Schematic representation of *S. cerevisiae* Snf1-pKD. The N- and C-lobes are coloured in light pink and pale green, respectively, and the activation segment is in orange. **f**, Superposition of the structures of *S. pombe* KD-AID, *S. cerevisiae* Snf1-pKD and Snf1-KD (PDB accession code 2FH9, in grey). In the close-up view on the right, corresponding KD-AID interface residues from Snf1-pKD are highlighted in cyan and labelled in parenthesis.

kinase domain and rat proteins. These data indicate that the dimers observed in our crystals are probably crystallization artefacts. The structural analyses hereafter were on the basis of monomer A of *S. cerevisiae* Snf1-pKD, and monomer B of *S. pombe* KD-AID, with lower average temperature factors (Supplementary Table 1).

Although no homologue structure was found for KD-AID in the PDB using the Dali server, the kinase domain and AID separately show high similarity to the kinase domain and ubiquitin-associated (UBA) domain of the AMPK-related protein kinases MARKs[24]. Both the AID of AMPK and the UBA domain of MARK fold into a non-canonical UBA conformation, yet they differ in both their interaction mode and regulatory function (Fig. 2 and Supplementary Fig. 6). In MARK, the UBA domain binds exclusively to the N-lobe of its kinase domain, and is required for its phosphorylation[24,25]. In contrast, the AID binds to the hinge region, opposite to the catalytic cleft of the kinase domain, and is not essential for the phosphorylation/activation of AMPK[15] (Fig. 1). In KD-AID, both the N- and C-lobes of the kinase domain are engaged with the AID, resulting in the burial of ~1,500 Å² exposed surface area. The predominantly hydrophobic interactions mainly involve four helices: α1 and α3 from the AID and αC and αE from the kinase domain. In particular, the invariant Leu 341 on helix α3 penetrates into a groove lined by hydrophobic residues from both lobes of the kinase domain (Fig. 2c). In turn, highly conserved Leu 88 at the C terminus of helix αC nestles into a hydrophobic pocket on the AID, gated by Met 316 of helix α1 and Leu 342 of helix α3. As well as the hydrophobic contacts, the negatively charged Glu 344 on helix α3 of the AID interacts with positively charged Arg 149 from the C-lobe, which also hydrogen bonds to the linker connecting the AID with its kinase domain (Fig. 2d). The side chain of the next residue Asn 345 is in contact with Tyr 146 from the C-lobe and Arg 90 from the N-lobe. This interaction mode contrasts sharply with that observed in MARK, in which the UBA domain binds, mainly through its helix α3, onto the concave outer face of the N-lobe β-sheet of its kinase domain (Supplementary Fig. 6b). In view of the marked difference, the observed interaction mode between AMPK kinase domain and its AID is distinct from that in the predicted model for the human AMPK α1-subunit fragment[8].

Many kinases are regulated, in addition to phosphorylation on the activation segments, through protein–protein interactions involving their kinase domains[22,23,26]. We wondered how AID-binding retains the kinase domain in a low-activity conformation. A close-up view of the superposition of kinase domains from *S. pombe* KD-AID and Snf1-pKD shows that residues from helix αC of a phosphorylated kinase domain could introduce a steric clash with the side chains of the key interface residues on the AID. For instance, owing to the inward and downward movement of helix αC, the ambulant side chain of Leu 109 in Snf1-pKD (Leu 88 in *S. pombe* KD-AID) may block the hydrophobic pocket for AID binding (Fig. 1f). Conversely, AID binding probably constrains the mobility of helix αC, resulting in a 'relative open' kinase conformation. Thus, this result discloses a unique mechanism for modulating kinase activity by directly interacting with helices αC and αE, although the interaction mode of AID binding to the backside of its kinase domain is reminiscent of the intramolecular autoinhibition mode observed in ZAP70 (ref. 27).

To assess the importance of the aforementioned interactions, we generated a series of point mutations on the *S. pombe* KD-AID fragment and examined their effect on catalytic function (Fig. 3a). When the key hydrophobic residues on the AID interface were individually replaced by charged residues (L341D, L342D and M316E), their catalytic efficiencies were increased approximately tenfold, comparable to that of the wild-type kinase domain. Mutant E344K, engineered to disrupt the hydrophilic interactions, yielded a modest but marked increase in basal activity, whereas substitution of the next residue (N345A) had little effect. To determine the role of the short linker connecting the AID to the kinase domain, we mutated Arg 280 that anchors the linker via four hydrogen bonds (Fig. 2c). The R280A mutation had no effect on kinase activity. However, mutating Leu 312 at the hydrophobic core of the AID to charged Asp (L312D) led to marked enzymatic activation (Supplementary Fig. 6d), which indicates that integrity of the AID is also required for its inhibitory function. We also mutated two crucial residues on the kinase domain interface, L88A and R149E, and these kinase-domain mutants had decreased catalytic activities. Nevertheless, the corresponding mutants in the background of KD-AID had similar kinase activities,

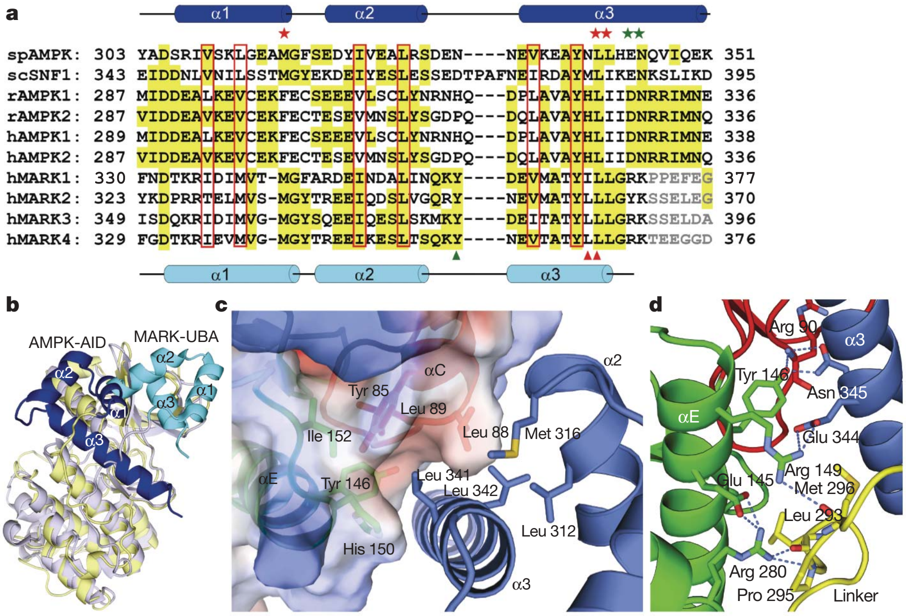


Figure 2 | The AID immobilizes the open conformation of kinase domain. **a**, Sequence alignment of AMPK-AIDs and MARK-UBAs. The hydrophobic core residues are boxed in red. Residues of AMPKs involving in hydrophobic and hydrophilic contacts with kinase domain are indicated by red and green asterisks, respectively, and those of MARKs by triangles. **b**, Structural comparison of *S. pombe* KD-AID with MARK1 KD-UBA (PDB accession code 2HAK). The kinase domain and the AID of AMPK are coloured in yellow and blue, respectively, and the kinase domain of MARK1 is in blue-white and the UBA domain is in cyan. **c**, van der Waals contacts at the KD-AID interface. The kinase domain is shown in surface representation, coloured according to electrostatic potential (positive, blue; negative, red). **d**, Hydrophilic interactions between the AID and the kinase domain.

indicating that these mutations also abolished AID inhibition (data not shown). These results indicate that the hydrophobic contacts between the kinase domain and the AID have a predominant role in controlling the conformational change between low- and high-activity forms of AMPK.

Because autoinhibition by the AID is universal to AMPK subfamily kinases, and as residues buried in the hydrophobic core of the AID are highly conserved (Figs 1 and 2), it is reasonable to postulate that rat

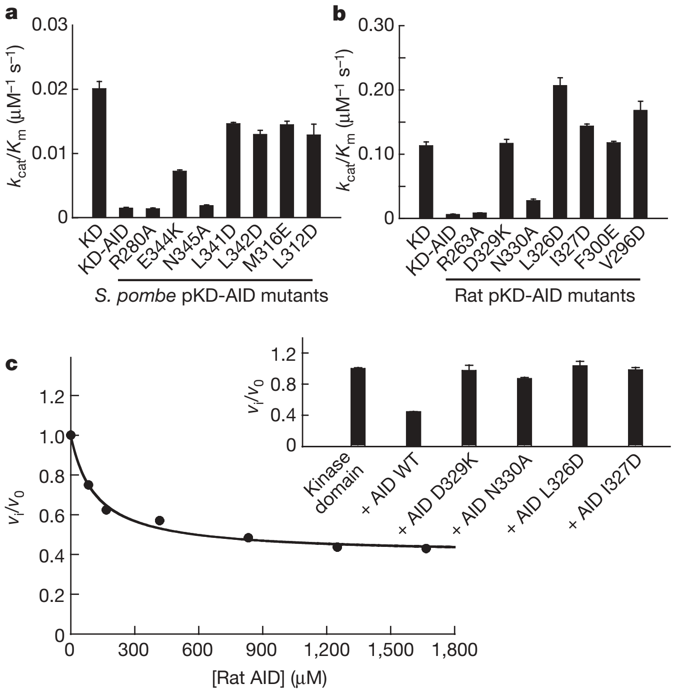


Figure 3 | Disruption of KD-AID interface abolishes AID inhibition. **a**, **b**, Comparison of $k_{cat}/K_{m(SAMS)}$ values of KD-AID mutants (mean and s.e.m., $n = 3$). The assay was performed in the presence of 1 mM ATP, 10 μM SAMS and either 500 nM *S. pombe* (**a**) or 100 nM rat (**b**) AMPK proteins. **c**, *Trans*-inhibition of the rat AID on the activity of its kinase domain. The assay was performed in the presence of 100 nM kinase domain, 20 μM SAMS, 1 mM ATP and various concentrations of the AID. The inset compares the inhibitory effects of wild-type AID and its mutants (mean and s.e.m., $n = 3$).

AMPK also has a functionally and structurally conserved autoinhibitory domain. We generated rat AMPK-α1 (also known as PRKAA1) mutants bearing mutations corresponding to those on the *S. pombe* KD-AID and examined their effects on kinase activity. As expected, in rat KD-AID mutants R263A and N330A there was little or no effect, whereas all others showed enhanced specific activities of more than 15-fold (Fig. 3b). To confirm the role of the AID further, we carried out a *trans*-inhibition assay using isolated rat AID, and the incomplete inhibition plateauing at ~50% provides evidence for the non-active-site binding of the AID (Fig. 3c and Supplementary Fig. 6e). Consistent with the effects of mutating the interface residues of the AID in *cis*-configuration, mutations in the isolated AID removed its *trans*-inhibition on the catalytic activity of kinase domain (Fig. 3c, inset).

As mammalian AMPK can be allosterically regulated by AMP binding to the γ-subunit, we next examined the effect of these mutations on the AMP activation of the rat AMPK holoenzyme. The wild-type heterotrimer is activated about twofold in the presence of 200 μM AMP, as reported[7] (Supplementary Fig. 7a). Mutant N330A, with a slightly higher initial rate, retained a certain AMP-dependence, whereas the intrinsic activities of the remaining mutants were increased independently of AMP (Fig. 4a and Supplementary Fig. 7b). These data suggest that the AID is central to the allosteric control of AMPK by AMP. Therefore, we consider that the inhibition of AMPK kinase activity is probably due to the AID interaction but not the proposed pseudosubstrate contribution from the γ-subunit[28]. Intriguingly, some of the mutants exhibited even higher activity than AMP-activated wild-type AMPK, which highlights the potential to develop new types of small compounds that activate AMPK by antagonizing the autoinhibition role of the AID, such as PT1 (ref. 29).

In addition to its allosteric effect, AMP was reported to markedly reduce the dephosphorylation rate of AMPK[6,30]. Incubation of wild-type AMPK with human protein phosphatase 2C-α (PP2C-α, also known as PPM1A) alone led to marked dephosphorylation of pThr 172, whereas the addition of AMP significantly slowed down the dephosphorylation rate (Fig. 4b and Supplementary Fig. 8). In contrast, the interface mutants drastically reduced the pThr 172-dephosphorylation rate and largely abolished the AMP-dependence. These results indicate that the AID interaction may facilitate the dephosphorylation, and that this effect can be reversed by disruption of the KD-AID interface through mutation or by AMP-binding to the γ-subunit. Together, these mutagenesis data suggest a new conformational switch model for AMPK regulation: AMP binding to the

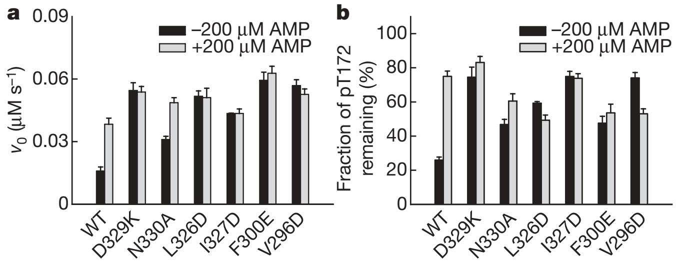


Figure 4 | Effect of KD-AID interface mutants on AMP-regulation of AMPK holoenzyme. a, Activity of AMPK heterotrimeric mutants independent of AMP change (mean and s.e.m., $n = 3$). The assay was performed in the presence of 2 nM AMPK, 1 mM ATP and 100 μM SAMS, with or without the addition of 200 μM AMP. **b**, Effects of mutations on AMP-protection of pThr 172 dephosphorylation (mean and s.e.m., $n = 3$). Phosphorylated AMPKs (7 μM) were treated with PP2Cα (0.6 μM) for 40 min at 25 °C, in the absence or presence of AMP (200 μM).

γ-subunit might transmit, by a yet unknown mechanism, onto the AID, alter the interaction between the AID and the kinase domain, and ultimately remove the effect of the AID on both kinase activation and pThr 172 dephosphorylation (Supplementary Fig. 9).

A long-standing paradox concerning the regulation of AMPK is that the α-subunit, in its isolated form, has little kinase activity and is apparently inhibited by the autoinhibitory sequence, and yet the AMPK holoenzyme is in a relatively active state[7]. To address this question, we further evaluated the role of the AID within heterotrimeric AMPK by detailed kinetic analyses. These analyses showed that the AID in the holoenzyme has a bona fide inhibiting role in the rate of phosphoryl transfer (catalytic constant, k_{cat}) as it does in the KD-AID, and that the catalytic activity and substrate binding affinity of AMPK are separately regulated by AMP binding and the assembly of β- and γ-subunits onto the α-subunit (Supplementary Table 2). Taken together, we have unravelled the structural basis of the autoinhibition of AMPK by the AID. The regulatory mode of AMPK by AMP provides a new mechanism by which autoinhibitory kinases are regulated by physiological signals.

METHODS SUMMARY

The kinase domain and KD-AID fragments of AMPKs from yeast and rat were subcloned and purified. The rat AMPK holoenzyme was expressed in a tricistronic vector. All mutants were generated by overlap PCR procedure. The purified proteins were stored at −80 °C and glycerol was added to stocks used for enzymatic assay. Crystals of *S. pombe* KD-AID and *S. cerevisiae* Snf1-pKD were grown by mixing proteins with ammonium sulphate and ammonium formate, respectively. The structures were solved by molecular replacement and additional residues (of the AID fragment) were manually located. The data processing and refinement statistics are summarized in Supplementary Table 1. AMPK proteins were phosphorylated by CaMKKβ, and the enzymatic activity of AMPK was determined by a coupled spectrophotometric assay. The kinetic parameters were obtained by fitting the experimental data to the Michaelis–Menten equation. The dephosphorylation of AMPK by PP2Cα was analysed with western blotting.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 14 December 2008; accepted 17 April 2009.
Published online 27 May 2009.

1. Carling, D. The AMP-activated protein kinase cascade—a unifying system for energy control. *Trends Biochem. Sci.* **29,** 18–24 (2004).
2. Kahn, B. B., Alquier, T., Carling, D. & Hardie, D. G. AMP-activated protein kinase: ancient energy gauge provides clues to modern understanding of metabolism. *Cell Metab.* **1,** 15–25 (2005).
3. Hardie, D. G. AMP-activated/SNF1 protein kinases: conserved guardians of cellular energy. *Nature Rev. Mol. Cell Biol.* **8,** 774–785 (2007).
4. Hardie, D. G. AMP-activated protein kinase as a drug target. *Annu. Rev. Pharmacol. Toxicol.* **47,** 185–210 (2007).
5. Yeh, L. A., Lee, K. H. & Kim, K. H. Regulation of rat liver acetyl-CoA carboxylase. Regulation of phosphorylation and inactivation of acetyl-CoA carboxylase by the adenylate energy charge. *J. Biol. Chem.* **255,** 2308–2314 (1980).
6. Davies, S. P., Helps, N. R., Cohen, P. T. & Hardie, D. G. 5′-AMP inhibits dephosphorylation, as well as promoting phosphorylation, of the AMP-activated protein kinase. Studies using bacterially expressed human protein phosphatase-2Cα and native bovine protein phosphatase-2A$_C$. *FEBS Lett.* **377,** 421–425 (1995).
7. Crute, B. E., Seefeld, K., Gamble, J., Kemp, B. E. & Witters, L. A. Functional domains of the α1 catalytic subunit of the AMP-activated protein kinase. *J. Biol. Chem.* **273,** 35347–35354 (1998).
8. Pang, T. *et al.* Conserved α-helix acts as autoinhibitory sequence in AMP-activated protein kinase α subunits. *J. Biol. Chem.* **282,** 495–506 (2007).
9. Hawley, S. A. *et al.* Characterization of the AMP-activated protein kinase kinase from rat liver and identification of threonine 172 as the major site at which it phosphorylates AMP-activated protein kinase. *J. Biol. Chem.* **271,** 27879–27887 (1996).
10. Scott, J. W. *et al.* CBS domains form energy-sensing modules whose binding of adenosine ligands is disrupted by disease mutations. *J. Clin. Invest.* **113,** 274–284 (2004).
11. Iseli, T. J. *et al.* AMP-activated protein kinase β subunit tethers α and γ subunits via its C-terminal sequence (186–270). *J. Biol. Chem.* **280,** 13395–13400 (2005).
12. McBride, A., Ghilagaber, S., Nikolaev, A. & Hardie, D. G. The glycogen-binding domain on the AMPK β subunit allows the kinase to act as a glycogen sensor. *Cell Metab.* **9,** 23–34 (2009).
13. Cool, B. *et al.* Identification and characterization of a small molecule AMPK activator that treats key components of type 2 diabetes and the metabolic syndrome. *Cell Metab.* **3,** 403–416 (2006).
14. Sanders, M. J. *et al.* Defining the mechanism of activation of AMP-activated protein kinase by the small molecule A-769662, a member of the thienopyridone family. *J. Biol. Chem.* **282,** 32539–32548 (2007).
15. Göransson, O. *et al.* Mechanism of action of A-769662, a valuable tool for activation of AMP-activated protein kinase. *J. Biol. Chem.* **282,** 32549–32560 (2007).
16. Scott, J. W. *et al.* Thienopyridone drugs are selective activators of AMP-activated protein kinase β1-containing complexes. *Chem. Biol.* **15,** 1220–1230 (2008).
17. Townley, R. & Shapiro, L. Crystal structures of the adenylate sensor from fission yeast AMP-activated protein kinase. *Science* **315,** 1726–1729 (2007).
18. Amodeo, G. A., Rudolph, M. J. & Tong, L. Crystal structure of the heterotrimer core of *Saccharomyces cerevisiae* AMPK homologue SNF1. *Nature* **449,** 492–495 (2007).
19. Xiao, B. *et al.* Structural basis for AMP binding to mammalian AMP-activated protein kinase. *Nature* **449,** 496–500 (2007).
20. Rudolph, M. J., Amodeo, G. A., Bai, Y. & Tong, L. Crystal structure of the protein kinase domain of yeast AMP-activated protein kinase Snf1. *Biochem. Biophys. Res. Commun.* **337,** 1224–1228 (2005).
21. Nayak, V. *et al.* Structure and dimerization of the kinase domain from yeast Snf1, a member of the Snf1/AMPK protein family. *Structure* **14,** 477–485 (2006).
22. Nolen, B., Taylor, S. & Ghosh, G. Regulation of protein kinases: controlling activity through activation segment conformation. *Mol. Cell* **15,** 661–675 (2004).
23. Pellicena, P. & Kuriyan, J. Protein–protein interactions in the allosteric regulation of protein kinases. *Curr. Opin. Struct. Biol.* **16,** 702–709 (2006).
24. Marx, A. *et al.* Structural variations in the catalytic and ubiquitin-associated domains of microtubule-associated protein/microtubule affinity regulating kinase (MARK) 1 and MARK2. *J. Biol. Chem.* **281,** 27586–27599 (2006).
25. Jaleel, M. *et al.* The ubiquitin-associated domain of AMPK-related kinases regulates conformation and LKB1-mediated phosphorylation and activation. *Biochem. J.* **394,** 545–555 (2006).
26. Huse, M. & Kuriyan, J. The conformational plasticity of protein kinases. *Cell* **109,** 275–282 (2002).
27. Deindl, S. *et al.* Structural basis for the inhibition of tyrosine kinase activity of ZAP-70. *Cell* **129,** 735–746 (2007).
28. Scott, J. W., Ross, F. A., Liu, J. K. & Hardie, D. G. Regulation of AMP-activated protein kinase by a pseudosubstrate sequence on the gamma subunit. *EMBO J.* **26,** 806–815 (2007).
29. Pang, T. *et al.* Small molecule antagonizes autoinhibition and activates AMP-activated protein kinase in cells. *J. Biol. Chem.* **283,** 16051–16060 (2008).
30. Sanders, M. J. *et al.* Investigating the mechanism for AMP activation of the AMP-activated protein kinase cascade. *Biochem. J.* **403,** 139–148 (2007).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank Y. Shi and S. C. Lin for critical discussions and reading of the manuscript; J. Chai and Y. Dong for help with data collection and processing; L. Gu and J. Wang for advice on structure determination. This work is supported in part by MOST grants 2006CB503900 and 2007CB914400, and by NSFC grants 30425005 and 30770476.

Author Contributions L.C., Z.-H.J. and L.-S.Z. designed, performed and analysed most of the experiments. Y.-Y.Z. and S.-T.X. provided technical assistance. Z.-X.W. contributed to discussions. J.-W.W. led the team and wrote the paper.

Author Information The structural and atomic coordinates have been deposited in the Protein Data Bank under accession codes 3H4J for *S. pombe* KD-AID and 3DAE for *S. cerevisiae* Snf1-pKD. Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to J.W.W. (jiaweiwu@mail.tsinghua.edu.cn).

METHODS

Constructs, mutagenesis and protein purification. The KD-AID (residues 25–351) and kinase domain (25–297) of *S. pombe* AMPK α-subunit were amplified by standard PCR procedure from complementary DNA of strain SPQ-01 (provided by J. Liang) and inserted into pET28b vector by NcoI/XhoI sites with C-terminal His_6-tag. The KD-AID (residues 41–398) and kinase domain (41–315) of *S. cerevisiae* AMPK were amplified from genomic DNA of *S. cerevisiae* S288C (Invitrogen) and inserted into pET21b vector. The KD-AID (residues 1–358), kinase domain (1–280) and AID (284–358) of rat α1-subunit (cDNA provided by P. Li) were also cloned into pET28b. The rat AMPK holoenzyme (α1β1γ1) was expressed in a tricistronic vector as reported[31]. All mutants were generated by overlap PCR procedure and subjected to DNA sequencing. The expression vector of CaMKKβ was provided by C. R. Mena. All proteins, overexpressed in *E. coli* BL21(DE3) cells at 20 °C, were first purified over Ni-NTA columns (Qiagen) and further polished by ion exchange and gel filtration chromatography (Source-15Q/15S and Superdex-200, GE Healthcare). The KD-AID fragments and kinase domains from *S. pombe*, *S. cerevisiae* and rat, phosphorylated or unphosphorylated, were stored at −80 °C and subjected to crystallization trials. Glycerol at a final concentration of 20% (v/v) was added into protein stocks used for enzymatic assay.

Crystallography. Crystals of *S. pombe* KD-AID were grown by the hanging-drop vapour diffusion method by mixing protein (~10 mg ml^{-1}) with an equal volume of reservoir solution containing 0.1 M sodium citrate, pH 5.6, 1.2 M ammonium sulphate. Crystals of *S. cerevisiae* Snf1-pKD were grown by mixing protein (~10 mg ml^{-1}) with an equal volume of reservoir solution containing 0.1 M Tris, pH 8.5, 4.0 M ammonium formate. Both crystals appeared after 2–3 days and grew to full size after 1 week. The crystals were cryoprotected in the reservoir solution supplemented with 20% glycerol and flash-frozen under cold nitrogen stream at 100K. The diffraction data set for KD-AID was collected at NW12 beam line at Photon Factory (Tsukuba, Japan) with an ADSC CCD detector, and that for Snf1-pKD collected at Beijing Synchrotron Radiation Facility with a CCD detector. Data were processed using HKL-2000 (ref. 32). The structures were solved by molecular replacement using Phaser[33] with one molecule from Snf1 kinase domain (PDB accession 3FAM) as search model[20]. Additional residues that were not included in the search model were manually located. Standard refinement was performed with the programs Phenix[34] and Coot[35]. The final model was refined to the R and R_{free} values of 21.8% and 25.7% for KD-AID and 22.7% and 26.6% for Snf1-pKD, respectively. The data processing and refinement statistics were summarized in Supplementary Table 1. All structural representations in this paper were prepared with Pymol (http://www.pymol.org).

Phosphorylation and activation of AMPK by CaMKKβ. To determine the phosphorylation and activation time of AMPK by CaMKKβ, the AMPK proteins were incubated in a phosphorylation buffer containing 50 mM Tris, pH 8.0, 2 mM DTT, 100 mM NaCl, 10 mM $MgCl_2$ and 1 mM ATP at 25 °C[36,37]. The reactions were initiated by adding CaMKKβ. Aliquots were withdrawn at indicated time intervals, and reactions were terminated by adding EDTA to a final concentration of 50 mM. The samples were then coupled to an AMPK kinase assay (see later) to directly examine the activation of AMPK by CaMKKβ, and a maximum activation was observed after 30 min of incubation (Supplementary Fig. 2). The identical samples were also subjected to western blot analysis to examine the phosphorylation state using an antibody against the phospho-Thr 172 of AMPK (Cell Signaling Technologies). In the following experiments, wild-type AMPK and all mutants (50 μM) were preincubated in the phosphorylation buffer with CaMKKβ (5 μM) and ATP (1 mM) for 1 h before the kinase activity assay.

Kinetic analysis of kinase activity. The enzymatic activity of AMPK was determined, with SAMS peptide as substrate, using a coupled spectrophotometric assay[38,39]. This assay couples the production of ADP with the oxidation of NADH by pyruvate kinase and lactate dehydrogenase (LDH). The standard assay was carried out at 25 °C in 1.8-ml reaction mixture containing 50 mM MOPS, pH 7.0, 100 mM NaCl, 0.1 mM EDTA, 10 mM $MgCl_2$, 0.2 mM NADH, 1.0 mM phospho(enol)pyruvate (PEP), 20 units ml^{-1} LDH, and 15 units ml^{-1} pyruvate kinase, and varying amounts of ATP, SAMS peptide and enzyme as indicated. Reactions were initiated by the addition of AMPK to the reaction mixture, and the enzymatic activity of AMPK was measured spectrophotometrically. Progress of the reaction was monitored continuously by following the formation of NAD^+ at 340 nm, on a PerkinElmer Lambda 45 spectrophotometer equipped with a magnetic stirrer in the cuvette holder. Initial rates were determined from the linear slope of progress curves. The concentrations of ADP formed in AMPK-catalysed reaction were determined using an extinction coefficient for NADH of 6220 cm^{-1} M^{-1}. The concentration of SAMS peptide was determined by turnover with the enzyme under conditions of limiting peptide at a fixed concentration of ATP. The kinetic parameters were obtained by fitting the experimental data to the Michaelis–Menten equation.

Dephosphorylation of AMPK by PP2Cα. The phosphorylated wild-type AMPK and mutants (7 μM) were incubated at 25 °C in a buffer containing 50 mM MOPS, pH 7.0, 100 mM NaCl, 0.1 mM EDTA and 2.5 mM $MgCl_2$, in the presence or absence of recombinant PP2Cα (0.6 μM) and AMP (200 μM). At indicated time intervals, aliquots were removed from the reservoir and the reaction was terminated by the addition of gel-loading buffer. Samples were resolved by SDS–PAGE and subjected to western blot analysis using a phospho-specific anti-AMPK pThr 172 antibody.

31. Neumann, D., Woods, A., Carling, D., Wallimann, T. & Schlattner, U. Mammalian AMP-activated protein kinase: functional, heterotrimeric complexes by co-expression of subunits in *Escherichia coli*. *Protein Expr. Purif.* **30**, 230–237 (2003).
32. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
33. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
34. Adams, P. D. *et al.* PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr. D* **58**, 1948–1954 (2002).
35. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).
36. Woods, A. *et al.* Identification of phosphorylation sites in AMP-activated protein kinase (AMPK) for upstream AMPK kinases and study of their roles by site-directed mutagenesis. *J. Biol. Chem.* **278**, 28434–28442 (2003).
37. Hawley, S. A. *et al.* Calmodulin-dependent protein kinase kinase-β is an alternative upstream kinase for AMP-activated protein kinase. *Cell Metab.* **2**, 9–19 (2005).
38. Roskoski, R. Jr. Assays of protein kinase. *Methods Enzymol.* **99**, 3–6 (1983).
39. Davies, S. P., Carling, D. & Hardie, D. G. Tissue distribution of the AMP-activated protein protein kinase, and lack of activation by cyclic-AMP-dependent protein kinase, studied using a specific and sensitive protein assay. *Eur. J. Biochem.* **186**, 123–128 (1989).

LETTERS

dUTP incorporation into genomic DNA is linked to transcription in yeast

Nayun Kim[1] & Sue Jinks-Robertson[1]

Highly activated transcription is associated with eukaryotic genome instability, resulting in increased rates of mitotic recombination and mutagenesis. The association between high transcription and genome stability is probably due to a variety of factors including an enhanced accumulation of DNA damage, transcription-associated supercoiling, collision between replication forks and the transcription machinery, and the persistence of RNA–DNA hybrids[1]. In the case of transcription-associated mutagenesis, we previously showed that there is a direct proportionality between the level of transcription and the mutation rate in the yeast *Saccharomyces cerevisiae*[2], and that the molecular nature of the mutations is affected by highly activated transcription[2,3]. Here we show that the accumulation of apurinic/apyrimidinic sites is greatly enhanced in highly transcribed yeast DNA. We further demonstrate that most apurinic/apyrimidinic sites in highly transcribed DNA are derived from the removal of uracil, the presence of which is linked to direct incorporation of dUTP in place of dTTP. These results show an unexpected relationship between transcription and the fidelity of DNA synthesis, and raise intriguing cell biological issues with regard to nucleotide pool compartmentalization.

The tetracycline/doxycycline-regulatable *TET-lys2*(ΔA746) frameshift allele detects net +1 mutations within an approximately 150-base-pair (bp) reversion window[4]. Although high levels of transcription stimulate a wide variety of mutations in this system[2], only the complex mutations, in which the selected frameshift is accompanied by one or more nearby base substitutions, are relevant to the experiments reported here. Because complex mutations are completely dependent on the activity of the translesion synthesis (TLS) DNA polymerase zeta (Polζ), they probably arise during lesion bypass and provide a unique molecular signature of this TLS polymerase[5,6].

Apurinic/apyrimidinic (AP) sites result either from spontaneous base loss or from the removal of a damaged base by the cognate DNA *N*-glycosylase, and their creation is the first step in the base excision repair (BER) pathway[7]. The BER pathway usually proceeds by AP endonuclease-catalysed cleavage of the sugar-phosphate backbone at the AP site, although some *N*-glycosylases have an associated AP lyase activity that also nicks the backbone. The subsequent steps in BER involve the removal of blocked ends, gap filling by a DNA polymerase and ligation of the remaining nick. More than 95% of the AP endonuclease activity in *S. cerevisiae* is due to the Apn1 protein[8]. When *APN1* was deleted in a strain containing the *TET-lys2*(ΔA746) allele, we found a 3.4-fold increase in the Lys^+ rate under high-transcription conditions (wild-type and *apn1*; Table 1), but no significant increase under low-transcription conditions (wild-type and doxycycline (Dox), and *apn1* and Dox; Table 1). Notably, this increase was accompanied by a 20-fold increase in the rate of complex mutations, with more than 50% occurring at a single site (Table 1 and Fig. 1). This site is named according to the mononucleotide run in which the selected frameshift occurs, and hence will be referred to as the 6A hotspot. When *REV3*, the gene encoding the catalytic subunit of Polζ, was deleted in addition to *APN1*, the Lys^+ rate was reduced 2.2-fold and no complex mutations were observed among 82 revertants sequenced. A similar result was obtained when *REV1* was deleted (data not shown).

Ntg1 and Ntg2 are functionally related proteins that possess AP lyase as well as *N*-glycosylase activity[9], and genetic data indicate that they can substitute for Apn1 in AP site processing[10,11]. Under high-transcription conditions, the overall reversion rate and the rate of complex mutations at the 6A hotspot were synergistically increased in an *apn1 ntg1 ntg2* triple mutant (10- and 18-fold, respectively) relative to the *apn1* single and *ntg1 ntg2* double mutant rates (Table 1 and Fig. 1). Of particular significance, 42 out of 43 complex mutations at the 6A hotspot in the triple mutant contained, in addition to the selected 6A-to-7A frameshift mutation, a T-to-G transversion at one of the two T residues immediately 3′ of the 6A run (highlighted in dark grey in Fig. 1; Table 2). As in the *apn1* single mutant, the increase in total Lys^+ rate and the rate of complex mutations at the 6A hotspot in the triple mutant were dependent on highly increased transcription and on the presence of Rev3 and Rev1.

The synergism observed between *apn1* and *ntg1 ntg2* suggests that a major lesion responsible for transcription-associated mutagenesis (TAM) in this system is an AP site. As well as Ntg1 and Ntg2, there are three other *N*-glycosylases in yeast: Ogg1, Mag1 and Ung1, which primarily remove 8-oxoguanine, methylated purines and uracil from DNA, respectively[7]. To determine whether any of these glycosylases might be responsible for generating the AP sites that accumulate under high-transcription conditions, each of the corresponding genes was deleted from the *apn1* background. The total reversion

Table 1 | Reversion of the *TET-lys2*(ΔA746) allele

Strain genotype	Lys^+ rate $\times 10^{-8}$ (95% CI)	Complex mutations at the 6A hotspot: Number/total	Rate ($\times 10^{-8}$)*
WT	4.28 (3.35–6.74)	1/117	0.037
WT + Dox	0.584 (0.438–0.612)	2/107	0.011
apn1	14.5 (12.4–18.1)	33/127	3.8
apn1 + Dox	0.646 (0.397–0.924)	4/89	0.029
apn1 rev3	6.54 (5.32–7.16)	0/82	<0.080
apn1 ogg1	22.7 (17.7–42.8)	24/93	5.9
apn1 mag1	20.9 (16.2–25.1)	31/93	7.0
apn1 ung1	7.79 (5.69–13.1)	8/92	0.68
ntg1 ntg2	5.19 (4.28–7.16)	2/93	0.11
apn1 ntg1 ntg2	150 (125–192)	43/94	69
apn1 ntg1 ntg2 + Dox	1.19 (0.949–1.32)	4/89	0.054
apn1 ntg1 ntg2 rev3	5.14 (2.65–10.2)	0/92	<0.057
apn1 ntg1 ntg2 rev1	4.54 (2.64–6.91)	1/88	0.052
apn1 ntg1 ntg2 ung1	8.15 (6.52–14.8)	4/89	0.37

Dox was added to repress transcription. CI, confidence interval.
* Rates were calculated by multiplying the total Lys^+ rate by the proportion of complex mutations at the 6A hotspot in the corresponding spectrum. When no such complex mutations were found, the rate was calculated assuming one event.

[1]Department of Molecular Genetics and Microbiology, Duke University Medical Center, Durham, North Carolina 27710, USA.

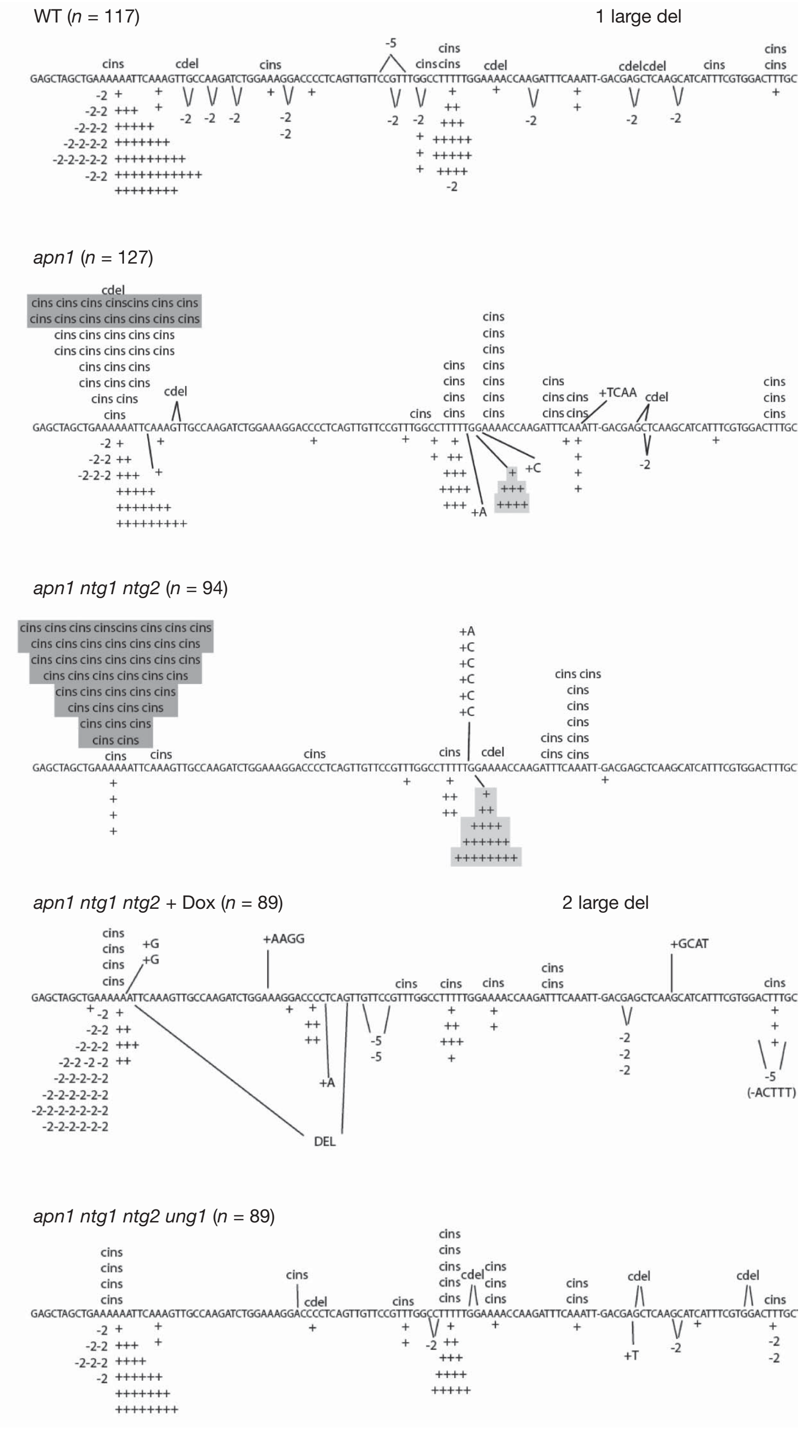


Figure 1 | *TET-lys2*(ΔA746) reversion spectra. Only the first 130 nucleotides of the 150-nucleotide reversion window are shown. The position of the A deletion that characterizes the allele is indicated by a dash within the sequence. Most simple insertions and deletions are indicated below the sequence by +, −2, −5, and so on; complex insertions and deletions (cins and cdel, respectively) are indicated above the sequence. Complex mutations at the 6A hotspot that are associated with a T-to-G transversion are highlighted in dark grey; insertions of G into a 2G run are highlighted in light grey. *n* denotes the number of independent revertants sequenced.

rate and the rate of complex mutations at the 6A hotspot were reduced only in the *apn1 ung1* double mutant. This effect was much more pronounced when *UNG1* was deleted in the *apn1 ntg1 ntg2* background, with the overall reversion rate decreasing more than 10-fold, and the rate of complex mutations at the 6A hotspot decreasing more than 100-fold (Table 1).

The marked Ung1-dependence of TAM in BER-deficient (*apn1* and *apn1 ntg1 ntg2*) strains demonstrates that uracil removal is the main source of AP sites in highly transcribed DNA. There are two known sources of uracil in DNA: spontaneous deamination of cytosine to uracil, and direct incorporation of dUTP during DNA synthesis. The incorporation of dUTP into genomic DNA is normally suppressed by the essential enzyme Dut1, which converts dUTP to dUMP[12]. To examine whether the incorporation of dUTP is responsible for TAM in the *apn1 ntg1 ntg2* mutant, the *DUT1*-encoded dUTPase was over-expressed by transforming mutant strains with a high-copy plasmid containing galactose-inducible *DUT1* (ref. 13). Transcription-associated and spontaneous mutagenesis were monitored in parallel by measuring *TET-lys2*(ΔA746) reversion and forward mutation at the *CAN1* locus, respectively. Although Dut1 overexpression had

little effect on the *CAN1* forward mutation rate, it was associated with a tenfold reduction in the *TET-lys2*(ΔA746) reversion rate (Supplementary Table 2), as well as a proportional decrease in complex insertions at the 6A hotspot (data not shown). Most of the uracil present in highly transcribed DNA is thus derived from direct dUTP incorporation rather than from cytosine deamination.

Previous analyses of complex mutation hotspots led us to suggest that these events result from misincorporation opposite a discrete lesion, followed by slippage in a short homopolymer run[5]. The most relevant features of the misincorporation-slippage model are that the site of the most frequent hotspot-associated base substitution marks the site of the initiating lesion, and that the nature of the most frequent base substitution reflects the identity of the lesion. The distinctive base substitution pattern at the 6A hotspot (AAAAAATT to AAAAAAAgT or AAAAAAATg; Table 2) is also more consistent with direct dUTP incorporation than with cytosine deamination, which would specifically result in base substitutions at cytosines. We suggest that dUTP is incorporated into the nontranscribed strand of the mutational target in place of one of the thymines immediately downstream of the 6A run (Fig. 2). After removal of the uracil by Ung1, specific insertion of dCTP during Polζ-assisted bypass of the resulting AP site will generate a T-to-G transversion. If bypass is followed by slippage in the 6A run that immediately follows, a complex mutation of the type observed will result. In agreement with the specificity of AP site bypass proposed here, it should be noted that a reduction in Dut1 activity has been associated with an Ung1-dependent increase in T-to-G transversions[13]. Although it is formally possible that the complex events at the 6A hotspot reflect a unique structural feature rather than a more general, transcription-associated increase in dUTP incorporation, the presence of a further hotspot in the *apn1* single and *apn1 ntg1 ntg2* triple mutants (TTTTTGG to TTTTTGGG; highlighted in light grey in Fig. 1) makes this unlikely. Of particular importance, the second hotspot has the same genetic requirements as the 6A hotspot (Supplementary Table 3), and can also be explained as a complex mutation resulting from cytosine insertion during AP site bypass coupled with slippage in a short homopolymer run (Supplementary Fig. 1).

Given the highly conserved nature of DNA structure and basic metabolic processes, it seems likely that TAM will be a general phenomenon affecting mutation accumulation in all organisms. The significance of the results reported here is that they directly link the availability of dUTP for DNA synthesis to the process of transcription. One intriguing possibility is that a locally high concentration of UTP is required to support efficient transcription and that this concomitantly increases the dUTP:dTTP ratio within the dNTP pool available to the replicative DNA polymerases. Such dUTP incorporation into transcriptionally active DNA could occur during normal DNA replication and/or during repair synthesis. These observations not only have broad implications for gene-specific mutagenesis and evolution, but also provide indirect evidence for nucleotide pool compartmentalization within the nucleus.

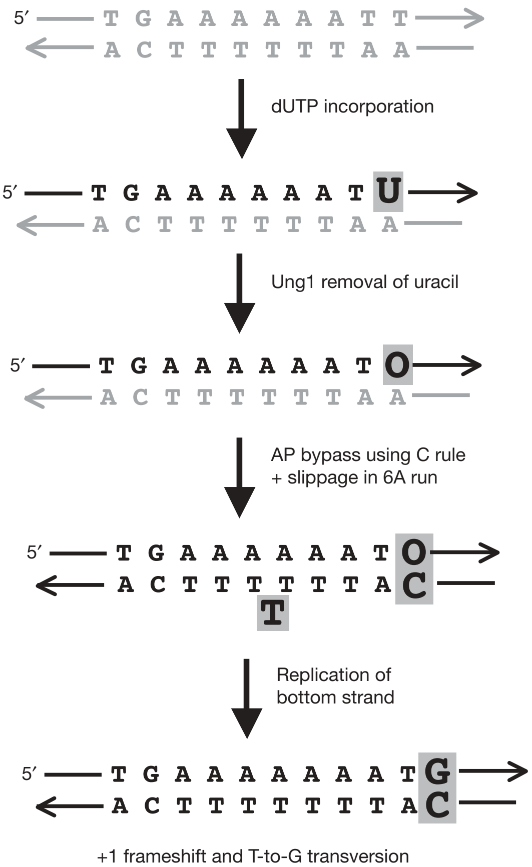


Figure 2 | Model for complex insertions at the 6A hotspot. Black lines and letters correspond to newly synthesized DNA. O denotes AP site.

Table 2 | Complex mutations at the 6A hotspot (AGCTGAAAAAATTC)

Sequence	Number of events in mutant background	
	apn1	*apn1 ntg1 ntg2*
AGCTGAAAAAAATgC	10	26
AGCTGAAAAAAAgTC	4	16
AGCTGCAAAAAATTC	4	0
AGCTGACAAAAATTC	5	0
AGCTGAACAAAATTC	2	0
AGCTcAAAAAAATTC	3	0
AGCTaAAAAAAATTC	2	0
AGCaGAAAAAAATTC	1	0
AGCTGAAAAAAAaTC	1	0
AGCTGAAAAgTC	1	0
AGCTGAAAAAAAAAATgC	0	1
Total	33	43

Insertions are underlined, base substitutions are in lowercase.

METHODS SUMMARY

The wild-type strain containing the *TET-lys2*(ΔA746) reporter on chromosome III near *ARS306* was previously described[2]. In all strains used here, transcription occurs in the same direction as replication fork movement. A PvuII fragment of p424-GAL1-DUT1 (ref. 13) was cloned into PvuII-digested pRS426 (ref. 14) to create p426-GAL1-DUT1. A complete strain list is given in Supplementary Table 1.

For mutation rates and spectra, 1-ml cultures (YEP supplemented with 2% glycerol plus 2% ethanol) with or without added doxycycline hyclate (2 $\mu g\,ml^{-1}$; Sigma) were inoculated with ~250,000 cells grown in the same medium. After 3 days, Lys^{+} revertants were selected on synthetic 2% dextrose plates deficient in lysine. Mutation rates were determined, and the median and 95% confidence intervals were calculated, as described previously[15] using data from 10 to 24 cultures. Sequencing of a PCR fragment amplified from independent revertants was done by the High Throughput Genomics Unit at the University of Washington.

Received 28 January; accepted 30 March 2009.
Published online 17 May 2009.

1. Aguilera, A. The connection between transcription and genomic instability. *EMBO J.* 21, 195–201 (2002).
2. Kim, N., Abdulovic, A. L., Gealy, R., Lippert, M. J. & Jinks-Robertson, S. Transcription-associated mutagenesis in yeast is directly proportional to the level of gene expression and influenced by the direction of DNA replication. *DNA Repair (Amst.)* 6, 1285–1296 (2007).
3. Lippert, M. J., Freedman, J. A., Barber, M. A. & Jinks-Robertson, S. Identification of a distinctive mutation spectrum associated with high levels of transcription in yeast. *Mol. Cell. Biol.* 24, 4801–4809 (2004).
4. Harfe, B. D. & Jinks-Robertson, S. Removal of frameshift intermediates by mismatch repair proteins in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* 19, 4766–4773 (1999).
5. Harfe, B. D. & Jinks-Robertson, S. DNA polymerase ζ introduces multiple mutations when bypassing spontaneous DNA damage in *Saccharomyces cerevisiae*. *Mol. Cell* 6, 1491–1499 (2000).

6. Minesinger, B. K. & Jinks-Robertson, S. Roles of RAD6 epistasis group members in spontaneous Polζ-dependent translesion synthesis in *Saccharomyces cerevisiae. Genetics* **169,** 1939–1955 (2005).
7. Boiteux, S. & Guillet, M. Abasic sites in DNA: repair and biological consequences in *Saccharomyces cerevisiae. DNA Repair (Amst.)* **3,** 1–12 (2004).
8. Popoff, S. C., Spira, A. S., Johnson, A. W. & Demple, B. The yeast structural gene (*APN1*) for the major apurinic endonuclease: homology to *E. coli* endonucelase IV. *Proc. Natl Acad. Sci. USA* **87,** 4193–4197 (1985).
9. Senturker, S. *et al.* Substrate specificities of Ntg1 and Ntg2 proteins of *Saccharomyces cerevisiae* for oxidized DNA bases are not identical. *Nucleic Acids Res.* **26,** 5270–5276 (1998).
10. Swanson, R. L., Morey, N. J., Doetsch, P. W. & Jinks-Robertson, S. Overlapping specificities of base excision repair, nucleotide excision repair, recombination, and translesion synthesis pathways for DNA base damage in *Saccharomyces cerevisiae. Mol. Cell. Biol.* **19,** 2929–2935 (1999).
11. You, H. J. *et al. Saccharomyces cerevisiae* Ntg1p and Ntg2p: broad specificity N-glycosylases for the repair of oxidative DNA damage in the nucleus and mitochondria. *Biochemistry* **38,** 11298–11306 (1999).
12. Gadsden, M. H., McIntosh, E. M., Game, J. C., Wilson, P. J. & Haynes, R. H. dUTP pyrophosphatase is an essential enzyme in *Saccharomyces cerevisiae. EMBO J.* **12,** 4425–4431 (1993).
13. Guillet, M., Van Der Kemp, P. A. & Boiteux, S. dUTPase activity is critical to maintain genetic stability in *Saccharomyces cerevisiae. Nucleic Acids Res.* **34,** 2056–2066 (2006).
14. Christianson, T. W., Sikorski, R. S., Dante, M., Shero, J. H. & Hieter, P. Multifunctional yeast high-copy-number shuttle vectors. *Gene* **110,** 119–122 (1992).
15. Spell, R. M. & Jinks-Robertson, S. in *Genetic Recombination: Reviews and Protocols* Vol. 262 (ed, Waldman, A. S.) 3–12 (Humana, 2004).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank S. Boiteux for providing the p424-GAL-DUT1 plasmid. We thank members of the laboratory and T. Petes for discussions and comments on the manuscript. This work was supported by a grant from the National Institutes of Health (R01 GM038464).

Author Contributions N.K. and S.J.-R. jointly designed experiments and wrote the manuscript. All experiments were performed by N.K.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to S.-J.R. (sue.robertson@duke.edu).

ERRATUM

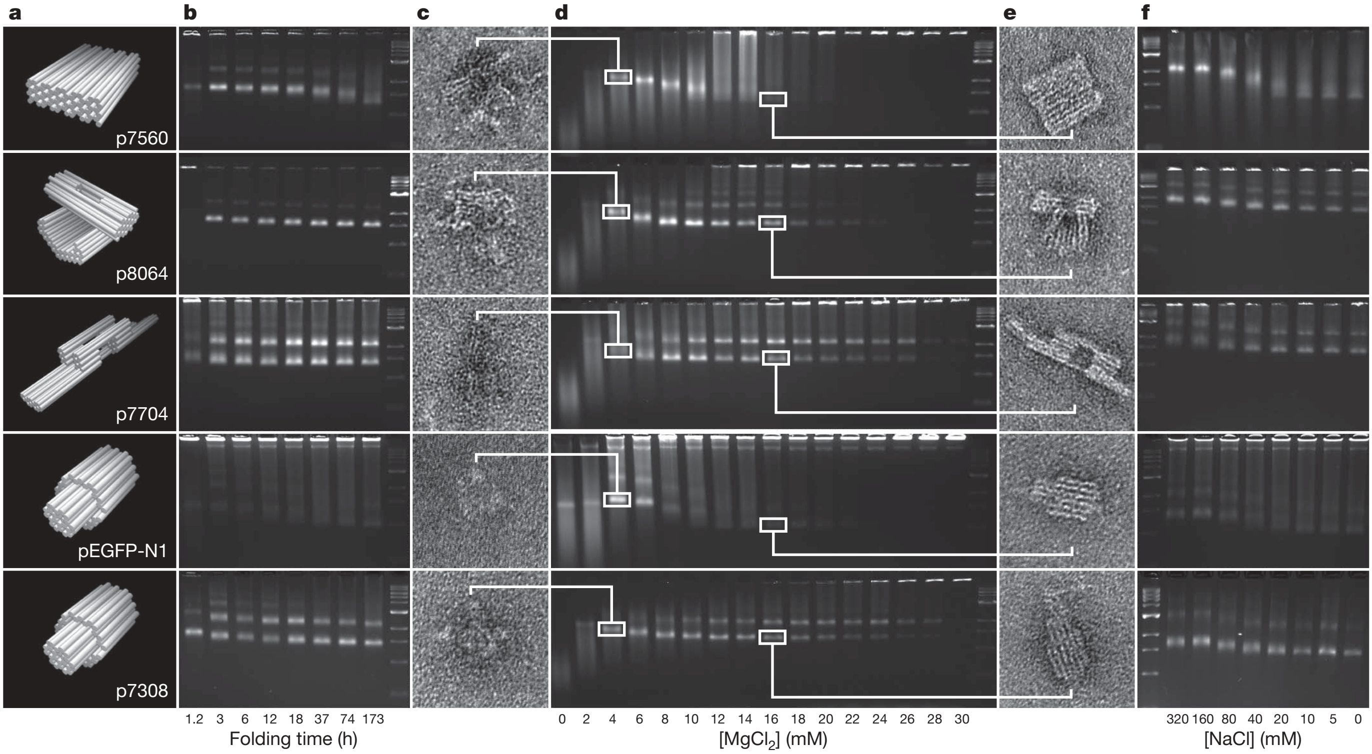
doi:10.1038/nature08165

Self-assembly of DNA into nanoscale three-dimensional shapes

Shawn M. Douglas, Hendrik Dietz, Tim Liedl, Björn Högberg, Franziska Graf & William M. Shih

Nature 459, 414–418 (2009)

In this Letter, Figure 3 was printed incorrectly. The corrected figure is presented below.



CORRIGENDUM

doi:10.1038/nature08166

The missing memristor found

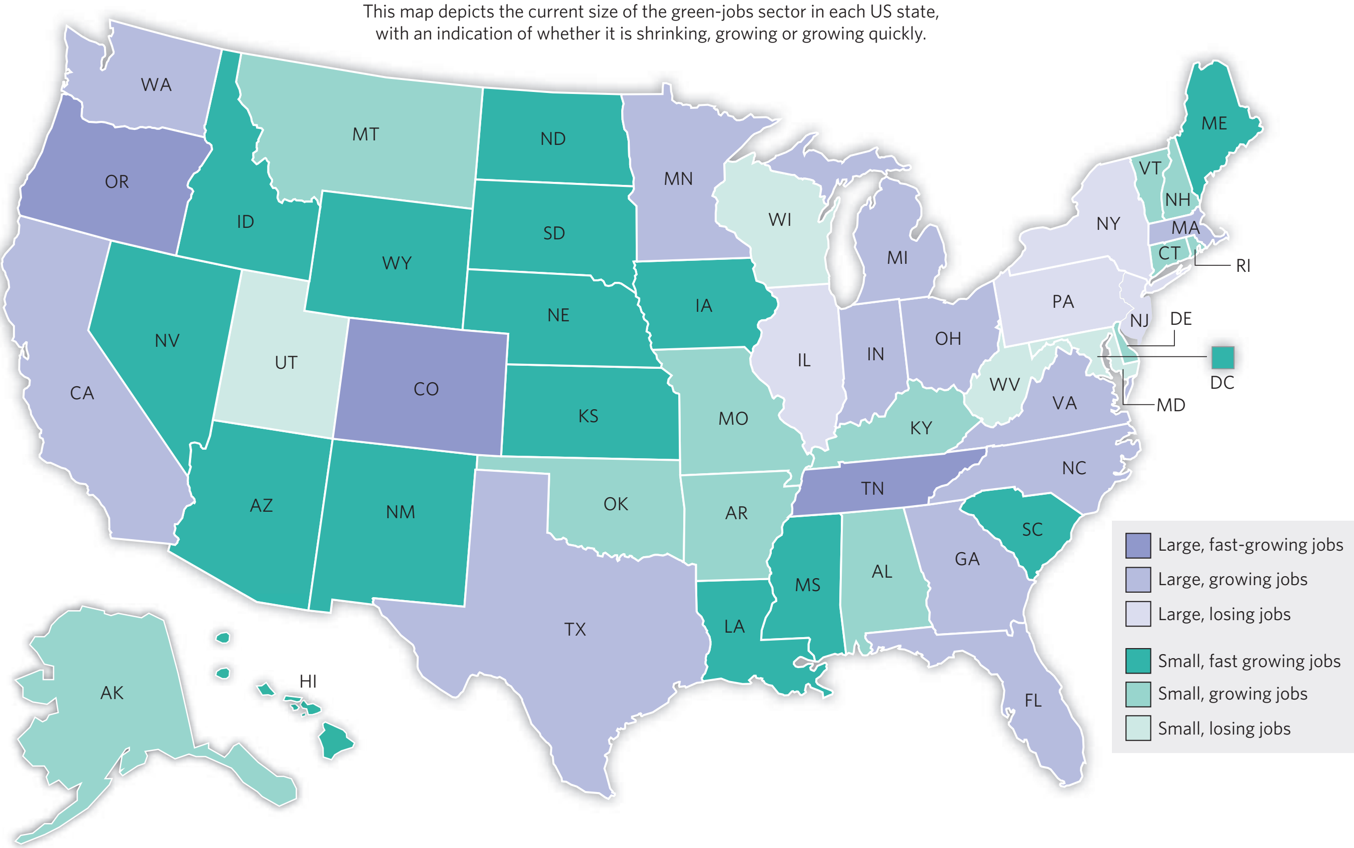
Dmitri B. Strukov, Gregory S. Snider, Duncan R. Stewart & R. Stanley Williams

Nature 453, 80–83 (2008)

In Fig. 2a of this Letter, the resistance for the right-hand element of the bottom circuit should be $R_{\mathrm{OFF}}(1 - w/D)$, instead of the shown value $R_{\mathrm{OFF}}w/D$. Also, the correct value of window function should be $w(D - W)/D^2$, rather than $w(1 - W)/D^2$. All simulation results presented in the paper use the correct formulae.

GREEN JOBS GROWTH

This map depicts the current size of the green-jobs sector in each US state, with an indication of whether it is shrinking, growing or growing quickly.



SOURCE: PEW CHARITABLE TRUSTS

NEWS

Green jobs growing, but destroying others?

Clean-energy research and engineering posts could be poised for a growth spurt in the United States if a predicted 'green-job' explosion becomes reality. But some critics suggest that green jobs — those with a role in reducing waste and pollution and benefiting the environment — are replacing other jobs and are costly to create.

The Clean Energy Economy, released on 10 June by the Pew Charitable Trusts, based in Washington DC, and *Climate 2030: A National Blueprint for a Clean Energy Economy*, released last month by the Union of Concerned Scientists in Cambridge, Massachusetts, are enthusiastic about the effect of clean energy on the US economy and about the potential for job growth. A related document from the Union of Concerned Scientists, *Clean Power, Green Jobs,* predicts that some 297,000 new green jobs will be created in sectors such as agriculture, forestry, manufacturing and construction by 2025. And the Pew report counted 770,000 existing green jobs in the United States as of 2007.

The Pew report, which claims to be the first analysis to count actual jobs, business and investments for all 50 states and the District of Columbia, notes that venture-capital investments in clean technology plunged 48% in the first quarter of 2009 compared with the same period the year before, but points out that that's still better than the 61% drop seen across all sectors. Clean-technology growth has varied widely from state to state, the report found. Nineteen states had more than the 2007 national average of 15,106 clean-energy jobs (referred to as 'large' states on map, below), and 18 'fast-growing' states had average annual growth between 1998 and 2007 that was above the national average of 1.9%.

Manufacturing gains

Representatives from both organizations agree that green-job growth is most likely to occur in the manufacturing and construction sectors, although they predict expansion in science and engineering research positions as well. Kil Huh, project director of research at the Pew Center on the States in Washington DC and lead researcher on the report, cannot estimate how many of the green jobs in the report are in science or science research. But he says that clean energy, energy efficiency and environmentally friendly production are magnets for venture capital and federal fiscal-stimulus investment, which, he predicts, will generate new research positions.

But economist Roger Meiners, a senior fellow with the Property & Environment Research Center, an environmental think tank based in Bozeman, Montana that in May published *7 Myths About Green Jobs*, says that green jobs actually cost the economy. The report says that in Spain, for example, each green job created has destroyed 2.2 existing jobs in other sectors.

The report says green-job outlays take resources from other sectors, raise energy prices, and encourage companies to move production facilities to lower-cost nations.

Meiners says building and construction, not research, is the focus of nearly all green jobs. "Most federal funds are designed to force construction of wind and sun technology," says Meiners. But Jeff Deyette, an energy analyst at the Union of Concerned Scientists, points to a clean-energy bill that is currently under congressional review. If passed, billions of dollars in federal funding will be directed to science research in clean energy and clean technology over the next several decades. Research targets such as wave power and hydrokinetics, nanotechnology and photovoltaic-cell technology are likely to receive federal funding, he says. ■

Karen Kaplan

Q&A

A long-time executive with Merck & Co, **Jeffrey Sturchio** recently took over as president and chief executive of the Global Health Council.

Did serendipity play a part in your career path?
Yes. I had no experience with international companies when I started at Merck as a company historian and archivist. It was a leap in the dark but also the opportunity to learn. From my initial role I was able to move into working with science policy. If I had pursued a master's degree in public health right after college, I'm not sure I would have been able to have as much impact.

What is the biggest challenge you face at the Global Health Council?
There are relatively few questions in global health for which we don't already have answers. We know how to prevent HIV transmission from mother to infant and what the appropriate treatments are for malaria. The real challenge is making sure that the people who know the answers are brought together with those who can implement them. Our goal is to organize the diverse expertise from all corners of the global-health communities.

What career accomplishment are you most proud of?
The role I had in helping millions of people get access to essential medicines. At Merck, we developed a programme that enabled hundreds of thousands of people in Africa and other parts of the developing world to obtain access to life-saving HIV medicines. We also had a programme to help 100 million people each year get free treatment to prevent river blindness and lymphatic filariasis.

What is the secret to your success?
I listen to people. It's no more complicated than that. If you want to solve a given problem, seek out experts and community members, listen carefully to what motivates them and what resources they need. Then find ways to establish common ground and bring together complementary skills. You don't have to agree on everything with your partners, but you can work towards a common goal if it is mutually beneficial.

Where do you find inspiration?
By getting out in the field and seeing health-care providers on the front lines of global health. It is amazing to watch people providing health care in refugee camps in the middle of nowhere with just basic materials, a shoestring budget and a lot of heart. The commitment and passion of these front-line workers — and the hope they bring — is deeply inspiring.

Have current global-health efforts made an impact?
Yes. The attention and resources flowing from global-health communities are creating tremendous excitement. We hope that the prevention of maternal–fetal HIV transmission will soon increase from 33% to 80% once the necessary resources are put in place. ■

Interview by Virginia Gewin

POSTDOC JOURNAL

Julia, you have to hustle

Babies are masters at causing chaos — and there's not a thing I can do about it

To help pay my way through graduate school, I worked in a women's clothing shop. During my sales training, the boutique owner pointed her finger at my chest. "Julia, you have to hustle!" she insisted. "And to hustle, you have to focus."

Her message — that to succeed I must stay sharp and on task — has stayed with me. Her voice resonates with me this month as I find myself haemorrhaging work hours thanks to childcare responsibilities.

The twice-daily day-care commute eats up precious work time and shortens my workday. My son's regular mingling with other children has also exposed him to a host of germs and necessitated my taking time off work for doctor's visits and at-home nursing care. I've been prey to some of these day-care 'superbugs', too. All of this has wiped out my productivity.

So where is my focus, other than on how inefficient I feel at work these days? Without a doubt, my attention and energy centre on my son's proper care and development. But the rest of my focus is splattered like baby food on the ceiling.

With a multitude of ongoing lab projects and papers that, to my impatient self, advance as if through thickest jelly, I need to regroup, buckle down and hustle. Otherwise, my focus will be on what I want to do rather than on my actual successes and what I can reasonably achieve. ■

Julia Boughner is a postdoc in evolutionary developmental biology at the University of Calgary, Canada.

IN BRIEF

US visa policy under fire

A statement from 31 academic and research associations in science and engineering urges the US federal government to streamline its visa process for international scientists, scholars and students. The 10 June statement, sent to several federal agencies, also called for a panel to be formed to review all federal visa policies and procedures enacted after the terrorist attacks of 11 September 2001. The group acknowledged improvements to the visa process, but said that delays are increasing. Its recommendations include a 30-day review for short-term applications from qualified researchers, scientists and graduate students and shorter processing for prominent repeat visitors.

The best places to work

Biotech firms Genentech and Monsanto, along with the University of Pennsylvania, were among the top five employers in an online poll ranking the 100 best places to work in the United States for information-technology professionals.

The poll, conducted by *Computerworld* magazine, ranks the 100 organizations on the basis of such criteria as training, job security, diversity and benefits. Genentech, based in South San Francisco, California, won kudos for its benefits plan and leadership programmes; Monsanto of St Louis, Missouri, gained a high rank for its flexible work hours and telecommuting policy; and the University of Pennsylvania in Philadelphia won praise for its benefits, diversity and secure-campus practices. Now in its 16th year, the survey collected responses from more than 27,000 employees across the United States.

Student gender bias

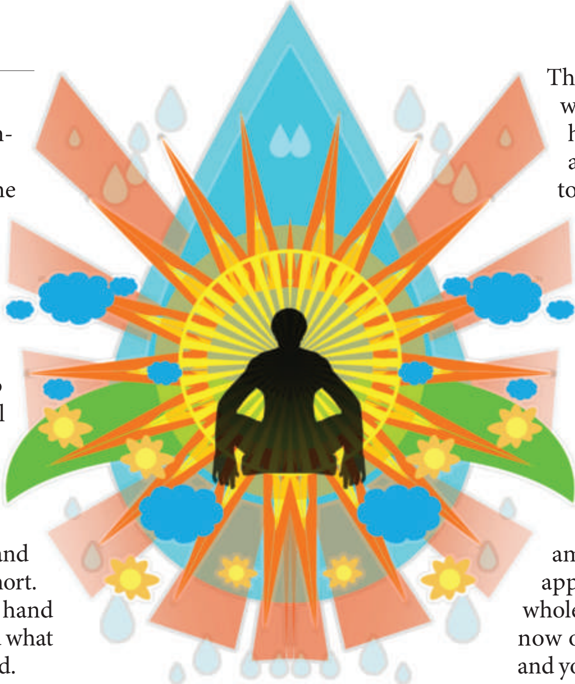
In the United Kingdom, more women than men are gaining admission to university, and they are more likely to stay once they get there, according to a study by the Higher Education Policy Institute. The Oxford-based non-profit group's report, *Male and Female Participation and Progression in Higher Education*, found that almost half of all women aged 17–30 attend college full- or part-time, compared with just over one-third of men of the same ages. About 8% of men aged 17–20 drop out, compared with 6.5% of women the same age, and 17% of men aged 21–30 drop out, compared with 12.3% of women the same age.

NATURE|Vol 459|25 June 2009

Codename: Phoenix

Here comes the rain.

Julian Tang



JACEY

It was blisteringly hot. The air shimmered above the desert sand.

The UN secretary-general and the country's leader stood just inside a large, battered, corrugated metal hangar. Their respective, air-conditioned limousines were parked farther inside.

They watched as a small propeller plane landed on the short airstrip outside. A young woman and a small boy, holding her hand, came carefully down the steps, out on to the hot sand.

The woman was dressed in a one-piece, white, sleeveless cotton dress and sandals. Her dark hair was cropped short. She looked sad and resigned. In the hand not holding onto the boy, she carried what looked like a large, square skateboard.

The boy seemed to be about five years old, with short brown hair, wearing a faded red-and-blue Spider-Man T-shirt with matching red shorts and sandals. In contrast to the woman, who must have been his mother, he looked around with the interest and fascination of any young child in new surroundings.

The leader looked sideways at his colleague.

"Just watch," the secretary-general said. "And whatever you see, whatever happens, do not interfere," he finished in a warning tone, glancing at the leader briefly to make sure he had been understood.

A quick nod from the leader satisfied him.

By now, the woman had put the skateboard on the sand between the plane and the hangar. She then returned to the plane and simply sat on the steps, watching the boy with some concern.

The boy took off his T-shirt, sandals and shorts and put them on the skateboard, making a kind of cushion for himself and sat down, cross-legged, now dressed only in his underwear. His manner suggested that this was a long-familiar routine. Then he just closed his eyes, rested his hands on his knees and faced up into the noon-day sun.

"This is crazy," whispered the leader. "He'll burn out there. We have to stop this."

He started forward, but the secretary-general put a firm hand on his arm to hold him back. "Remember what I said," he whispered back, fiercely. "Whatever you see, do *not* interfere — otherwise I will ask my assistant to restrain you." He nodded back towards his limo, where his assistant-driver-bodyguard had already got out and was now waiting by the car.

The leader nodded again in acceptance and stepped back into the hangar.

After about half an hour, the boy's skin began to turn pink. After another half hour, it became bright red.

The mother appeared to be weeping, but continued to watch the boy.

Almost imperceptibly, it seemed that the sky had become darker.

The leader looked up and saw that heavy rain clouds had gathered, as if from nowhere. Shortly after this, there were flashes of cloud-to-cloud lightning followed by distant thunder. The sun was now completely hidden. The boy's posture did not change.

Soon there was a torrential rain. Large water droplets began hammering down on the hangar roof creating a cacophonous din.

The rain did not stop for six days.

During this time, the woman stayed inside the plane with the pilot. They had both been through this before — the plane seemed to be well-stocked.

The leader had no provisions with him, but the secretary-general shared his with him. There was an old chemical toilet at the back of the hangar, but no shower.

The rain was so heavy that they could hardly see the boy, let alone the plane. The sand seemed to be saturated with water. New vegetation and wildlife had started appearing in the areas around the hangar that were visible to them. There were probably desert species taking sudden opportunistic advantage of the deluge.

The rain continued.

The leader thought that the air smelled different, fresher and greener, now with a cool breeze blowing in from across the desert.

After sleeping in his limo, he got out on the morning of the seventh day to find that the rain had stopped. He could hear birds singing. Looking out of the hangar in amazement, it seemed as if Eden had appeared just outside the door. The whole area, as far as the eye could see was now open grassland with small bushes and young trees scattered about. Running water could be heard coming from behind the hangar. Except for a small patch of sand underneath and around the skateboard and the plane, everything had been transformed into a green paradise.

On the skateboard were the remains of the boy.

In horror, the leader forgot himself and dashed out to the sandy patch and helplessly stared down at the blackened, burnt husk of what had been a healthy five-year old. He had a young son of about the same age.

He felt a gentle touch on his arm. "It's OK," the woman said quietly, looking kindly up into his face. "He'll be fine."

He looked down and saw that she had a wooden box in her hand. He then watched in disbelief as she used it to cover the body of the child, clipping it into position over the skateboard base then wheeling it to the back of the plane, over the new grass and up the cargo ramp.

The pilot returned to the cockpit and closed the cargo bay. Sitting inside, he saw the woman raise a hand to the UN secretary-general, who nodded in acknowledgement and gratitude. It seemed that they were old friends.

As the plane taxied away to prepare for take-off, the young boy, perfectly healthy again, appeared at the window and smiled and waved at him. ■

Julian Tang is a clinical/academic virologist. He would like to dedicate this story to his mum, who, like the mother in this story, was always standing by, ready to save her sons whenever required.